2020

Diversity As Impression Management

Edward Chang
University of Pennsylvania

Diversity As Impression Management

Abstract

In recent years, increasing scrutiny has been placed on groups and organizations and their levels of diversity. Because groups and organizations have incentives to avoid negative scrutiny and often engage in attempts to manage impressions around scrutinized behaviors, I propose that at least some organizations attend to diversity for impression management reasons. In Chapter 1, I use Monte Carlo simulations to provide evidence that S&P 1500 companies strategically manage the levels of gender diversity of their boards of directors for impression management reasons. I also show that scrutiny and visibility moderate these effects, consistent with an impression management explanation. In Chapter 2, I draw on the idea that organizations may attend to diversity for impression management reasons to explore a potential intervention–namely, having people make multiple hiring or selection decisions at once as opposed to making them one at a time–to increase gender diversity in organizations. Finally, in Chapter 3, I explore whether organizations engage in a form of impression management I name "diversity washing," whereby they mislead outsiders about their actual diversity levels or practices. I find evidence that members of certain demographic groups are systematically overrepresented in signals of diversity of some organizations, and this appears to be moderated by visibility. This dissertation highlights the importance of impression management in understanding contemporary diversity-related decisions in organizations.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Operations & Information Management

First Advisor

Katherine L. Milkman

Keywords

Diversity, Gender, Impression Management, Race

Subject Categories

Organizational Behavior and Theory

DIVERSITY AS IMPRESSION MANAGEMENT

Edward H. Chang

A DISSERTATION
in
Operations, Information and Decisions
For the Graduate Group in Managerial Science and Applied Economics
Presented to the Faculties of the University of Pennsylvania
in
Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2020

Supervisor of Dissertation

_____
Katherine L. Milkman, Professor of Operations, Information and Decisions

Graduate Group Chairperson

_____
Nancy R. Zhang, Professor of Statistics

Dissertation Committee
Maurice E. Schweitzer, Professor of Operations, Information and Decisions
Adam M. Grant, Professor of Management
Rebecca L. Schaumberg, Assistant Professor of Operations, Information and Decisions

ACKNOWLEDGMENT

I would like to express my utmost gratitude to my advisor, Katherine Milkman. Without her guidance and mentorship, none of this would have been possible. I cannot thank her enough for being my advisor.

I would also like to thank my dissertation committee–Maurice Schweitzer, Adam Grant, and Rebecca Schaumberg–and all of my collaborators on this and other research–Modupe Akinola, Dolly Chugh, Angela Duckworth, Dena Gromet, Erika Kirgios, Cade Massey, Aneesh Rai, and Robert Rebele. I have also received invaluable help and guidance from many other faculty members and students and would particularly like to thank Sigal Barsade, Bradford Bitterly, Celia Gaertig, Etan Green, Joshua Lewis, Mary-Hunter McDonnell, Lesley Meng, Robert Mislavsky, Alex Rees-Jones, Uri Simonsohn, and Sam Skowronek.

Finally, I am grateful for financial support from the Wharton Behavioral Lab, the Wharton Risk Center Russell Ackoff Doctoral Student Fellowship, and the Operations, Information and Decisions Department.

ABSTRACT

DIVERSITY AS IMPRESSION MANAGEMENT

Edward H. Chang

Katherine L. Milkman

In recent years, increasing scrutiny has been placed on groups and organizations and their levels of diversity. Because groups and organizations have incentives to avoid negative scrutiny and often engage in attempts to manage impressions around scrutinized behaviors, I propose that at least some organizations attend to diversity for impression management reasons. In Chapter 1, I use Monte Carlo simulations to provide evidence that S&P 1500 companies strategically manage the levels of gender diversity of their boards of directors for impression management reasons. I also show that scrutiny and visibility moderate these effects, consistent with an impression management explanation. In Chapter 2, I draw on the idea that organizations may attend to diversity for impression management reasons to explore a potential intervention–namely, having people make multiple hiring or selection decisions at once as opposed to making them one at a time–to increase gender diversity in organizations. Finally, in Chapter 3, I explore whether organizations engage in a form of impression management I name "diversity washing," whereby they mislead outsiders about their actual diversity levels or practices. I find evidence that members of certain demographic groups are systematically overrepresented in signals of diversity of some organizations, and this appears to be moderated by

visibility. This dissertation highlights the importance of impression management in understanding contemporary diversity-related decisions in organizations.

TABLE OF CONTENTS

INTRODUCTION

In recent years, negative scrutiny surrounding the lack of diversity in groups and organizations in myriad contexts has increased (Lee, 2017; Merchant, 2013; Ryan, 2016). Given that organizations have strong incentives to avoid negative scrutiny (Desai, 2011), past research has found that organizations engage in impression management tactics around scrutinized behaviors (Bolino, Kacmar, Turnley, & Gilstrap, 2008; Elsbach, Sutton, & Principe, 1998). Thus, if scrutiny surrounding diversity has increased, it is natural to predict that organizations may now be attending to diversity for impression management reasons.

By considering the possibility that organizations may demand diversity for impression management reasons, I integrate theoretical perspectives to help scholarship better understand diversity-related decisions in organizations. Taking impression management into account can help explain phenomena like why organizations cluster at the same levels of diversity and why individuals and organizations publicly proclaim support for diversity in spite of lack of progress towards diversifying (Fingerhut, 2018).

In Chapter 1, in collaboration with Katherine Milkman, Dolly Chugh, and Modupe Akinola, I provide evidence that impression management appears to drive diversity-related hiring decisions in some organizations. Using Monte Carlo simulations, we show that S&P 1500 boards are disproportionately likely to include exactly two women, a phenomenon we call "twokenism." Two women appears to be the minimum number of women needed to meet the descriptive social norm for diversity, suggesting

that S&P 1500 companies are striving for the bare minimum levels of gender diversity required to escape negative scrutiny. Consistent with this theorizing around scrutiny and impression management, we find that more visible companies are more likely to have exactly two women on their boards.

In Chapter 2, in collaboration with Erika Kirgios, Aneesh Rai, and Katherine Milkman, I explore a potential intervention to increase gender diversity in organizations that draws on the idea that people may attend to diversity for impression management concerns. Specifically, we propose that having people make multiple hiring or selection decisions at once–rather than making hiring or selection decisions one at a time–may lead them to select more gender diversity in candidates. Diversity may be more salient when selecting a group of candidates as opposed to when selecting an individual, and if people care about diversity for impression management reasons, they should be more motivated to avoid selecting an all-male group when diversity is more salient. In six preregistered experiments, we show that making multiple hiring or selection decisions at once does, in fact, increase the gender diversity of candidates selected compared to making identical hiring or selection decisions in isolation.

In Chapter 3, I document an impression management phenomenon I call "diversity washing," whereby companies mislead outsiders about their actual diversity levels or practices. In contrast to Chapter 1, which documents how impression management shapes actual levels of diversity in organizations, Chapter 3 explores impression management tactics organizations engage in that are completely decoupled

from actual diversity levels. Using Monte Carlo simulations, I show that Black people and women of all races are systematically overrepresented on technology company and law firm websites. Diversity washing appears to be exacerbated by increased visibility, consistent with past research on the influence of visibility on impression management behaviors.

Together, my dissertation explores how impression management may be a pervasive motivation driving diversity-related decisions in organizations. In particular, this work highlights the strength of impression management motives in guiding diversity-related decisions in organizations and suggests that research on diversity should consider impression management as a key motivation in theorizing and scholarship. In addition, by painting a fuller picture of what motivates organizations to consider and strive for diversity, I help provide insights into potential ways to increase diversity in organizations. Past interventions to increase diversity in organizations have found varying levels of success (Kalev, Dobbin, & Kelly, 2006). However, past efforts may have been unsuccessful because they assumed that we need to change people's biases and prejudices, which research has shown is challenging (Lai et al., 2016). If organizations are motivated to attend to diversity because of impression management, then that suggests additional levers we can use to influence diversity-related decisions within organizations that do not require changing people's conscious or unconscious biases.

References

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, *34*(6), 1080–1109.

Desai, V. M. (2011). Mass media and massive failures: Determining organizational efforts to defend field legitimacy following crises. *Academy of Management Journal*, *54*(2), 263–278.

Elsbach, K. D., Sutton, R. I., & Principe, K. E. (1998). Averting expected challenges through anticipatory impression management: A study of hospital billing. *Organization Science*, *9*(1), 68–86.

Fingerhut, H. (2018, June 14). Most in US say growing racial and ethnic diversity makes country better. *Pew Research Center*. https://www.pewresearch.org/fact-tank/2018/06/14/most-americans-express-positive-views-of-countrys-growing-racial-and-ethnic-diversity/

Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, *71*(4), 589–617.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., & Marshburn, C. K. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001.

Lee, J. C. (2017, January 13). Trump's Cabinet So Far Is More White and Male Than Any First Cabinet Since Reagan's. *The New York Times*. https://www.nytimes.com/interactive/2017/01/13/us/politics/trump-cabinet-women-minorities.html

Merchant, N. (2013, October 7). Viewpoint: Twitter's All-Male Board Spells Failure. *Time*. http://ideas.time.com/2013/10/07/viewpoint-twitters-all-male-board-spells-failure/

Ryan, P. (2016, February 2). #OscarsSoWhite controversy: What you need to know. *USA Today*. https://www.usatoday.com/story/life/movies/2016/02/02/oscars-academy-award-nominations-diversity/79645542/

CHAPTER 1. DIVERSITY THRESHOLDS: HOW SOCIAL NORMS, VISIBILITY, AND SCRUTINY RELATE TO GROUP COMPOSITION

Edward H. Chang, Katherine L. Milkman, Dolly Chugh, and Modupe Akinola

Published in *Academy of Management Journal* in 2019

ABSTRACT

Across a field study and four experiments, we examine how social norms and scrutiny affect decisions about adding members of underrepresented populations (e.g., women, racial minorities) to groups. When groups are scrutinized, we theorize that decision makers strive to match the diversity observed in peer groups due to impression management concerns, thereby conforming to the descriptive social norm. We examine this first in the context of U.S. corporate boards where firms face pressure to increase gender diversity. Analyses of S&P 1500 boards reveal that significantly more boards include exactly two women (the descriptive social norm) than would be expected by chance. This overrepresentation of two-women boards–a phenomenon we call "twokenism"–is more pronounced among more visible companies, consistent with our theorizing around impression management and scrutiny. Experimental data corroborate these findings and provide support for our theoretical mechanism: decision makers are discontinuously less likely to add a woman to a board once it includes two women (the social norm), and decision makers' likelihood of adding a woman or minority to a group is influenced by the descriptive social norms and scrutiny faced. Together, these findings

provide a new perspective on the persistent underrepresentation of women and minorities in organizations.

Link to *Online Supplement*:

https://osf.io/562yg/?view_only=1d8c31b6e5a94b0aa40ee90ac95f3f5a

INTRODUCTION

In recent years, many groups have faced negative scrutiny for their lack of diversity. For instance, the Academy of Motion Picture Arts and Sciences faced backlash in 2015 and in 2016 when all twenty actors nominated for Academy Awards in the lead and supporting acting categories were white. This sparked an #OscarsSoWhite meme and a plan to double female and minority membership in the Academy by 2020 (Ryan, 2016). When Twitter made an initial public offering with no women on its board of directors in 2013, the company faced an outpouring of negative media attention, with numerous outlets claiming that the lack of gender diversity would cause problems for the company (Merchant, 2013; Miller, 2013). And when Donald Trump announced the members of his presidential cabinet in 2017, the New York Times ran a front-page story tallying the women and racial minorities Trump's cabinet included and comparing its (lack of) diversity to all other modern U.S. administrations (Lee, 2017). These examples illustrate that when groups lack diversity, negative scrutiny–or critical attention paid to particular behaviors (Sutton & Galunic, 1996)–can ensue.

Little is known, however, about when a group's diversity will be judged negatively or how groups will respond to the possibility of negative scrutiny regarding their diversity. While scholarship has established that diversity is not perceived objectively, or equivalently, by all observers and in all contexts (Unzueta & Binning, 2010, 2012; Unzueta, Knowles, & Ho, 2012), it remains ambiguous as to when group members and those perceiving groups judge a group's diversity to be so insufficient as to

warrant action or attention. Further, although past work has established that organizations respond to reputational threats such as social movement boycotts (King, 2008; McDonnell & King, 2013), it is unclear how those responsible for group composition may behave when facing the threat of repercussions for displaying insufficient diversity. In this paper, we address these questions by analyzing a decade of data on the composition of U.S. corporate boards in the S&P 1500 and by conducting a series of supplemental experiments.

We propose that, to avoid facing negative scrutiny, those responsible for forming groups may seek safety in numbers by looking to the average behavior of others when setting implicit or explicit goals about the diversity of groups. Descriptive social norms–defined as the average observed behavior of individuals or groups in a population (Prentice & Miller, 1993)–have been shown to serve as reference points for behavior in a variety of contexts, setting expectations about what is appropriate and effective (Coffman, Featherstone, & Kessler, 2014; Goldstein, Cialdini, & Griskevicius, 2008; Nolan, Schultz, Cialdini, Goldstein, & Griskevicius, 2008), particularly in situations where appropriate behavior is ambiguous or uncertain (Festinger, 1954; Sherif, 1936). Decision makers and firms may thus look to relevant others to understand what the descriptive social norms for diversity are, and they may then imitate these levels of diversity, both because of the reputational threat associated with negative scrutiny and because of uncertainty about what adequate diversity entails (DiMaggio & Powell, 1983). This behavior should be even more prevalent among highly visible groups or

organizations because the negative consequences of failing to conform can be greater for high-profile groups (Gardberg & Fombrun, 2006). The actions of highly visible groups are more likely to be scrutinized in the first place (Chiu & Sharfman, 2011), and organizations generally respond more strongly to more visible threats (King, 2008).

We combine our theorizing about descriptive social norms, scrutiny, and visibility with past research on goal setting to make a novel prediction. Specifically, we predict that individuals responsible for group compositions will respond to pressures to diversify in a similar fashion, leading to an overabundance of groups with identical levels of diversity. Past research has shown goals–like the goal to match the diversity of peer groups–are often highly motivating (Locke & Latham, 2002), but individuals relax efforts to achieve desirable outcomes after reaching salient goal thresholds in many settings (Heath, Larrick, & Wu, 1999). This relaxing of effort has been shown to lead goal-seekers' performance to cluster around salient goal thresholds (Pope & Simonsohn, 2011). We predict this tendency will lead scrutinized groups to cluster around the social norm for diversity set by their peers. In other words, rather than continuing to increase diversity in response to external pressures (e.g., the threat of negative scrutiny), those with the power to shape group diversity should be less likely to increase the diversity of a group once the group has reached the descriptive social norm for diversity set by peers. This behavior will lead to improbably homogeneous diversity levels across groups.

We test our theorizing first in the context of U.S. corporate boards, a setting where firms face negative scrutiny for failing to include adequate gender diversity

(Merchant, 2013; Miller, 2013). Analyses of S&P 1500 boards reveal that significantly more boards include exactly two women (the descriptive social norm) than would be expected by chance, supporting our prediction that groups will respond to pressures to diversify in a similar fashion, leading to an overabundance of groups with identical levels of diversity at the descriptive social norm. This overrepresentation of two-women boards is more pronounced among more visible companies, consistent with our theorizing around impression management and scrutiny. In additional studies, we experimentally manipulate descriptive social norms, scrutiny, and visibility to show that each of these influences group diversity decisions as our theory predicts in groups besides corporate boards and when we examine social categories besides gender.

Our work provides a more complete understanding of diversity-related hiring decisions, telling us when women and racial minorities will be particularly attractive candidates for inclusion in groups and when groups will reduce their efforts to increase diversity. Further, rather than focusing only on individual-level or firm-level explanations for why women and racial minorities may or may not be added to groups, we highlight how external entities such as peers (who help shape descriptive social norms) and outsider scrutiny can shape group diversity decisions. By illuminating these critical factors that influence group diversity decisions, we provide theoretical guidance about potential new ways to improve diversity in organizations and practical guidance to help predict what levels of diversity we might expect to see in different contexts. Our research suggests that it may be helpful to increase scrutiny around diversity decisions and attempt

to make other social norms besides descriptive social norms salient to decision makers in order to increase the number of women and racial minorities selected into groups.

THEORY AND HYPOTHESES

Descriptive Social Norms

Descriptive social norms–defined as the average observed behavior of individuals or groups in a population (Prentice & Miller, 1993)–exert a potent influence on decisions. According to past research, descriptive social norms influence the behavior of individuals and groups for two primary reasons. First, they establish what is socially acceptable. Because following the norm means avoiding outlier status, individuals and groups can feel reassured that if existing norms are followed, social ostracism will not ensue (Schultz, Nolan, Cialdini, Goldstein, & Griskevicius, 2007). By following a descriptive social norm, individuals and groups essentially insulate themselves from the risk of being singled out because they are–by definition–doing what many of their peers are doing. Individuals, groups, and organizations that negatively deviate from any descriptive norm are much more likely to be singled out and face negative consequences (Ahmadjian & Robinson, 2001; Zavyalova, Pfarrer, Reger, & Shapiro, 2012).

Second, descriptive social norms contain information about what behaviors are likely to be effective or adaptive (Cialdini, 2007). If the majority of others have elected to partake in a specific action or behavior (making it the descriptive social norm), then that signals that the norm may be a wise course of action (e.g., if everyone else is using this brand of soap, it must be a good brand of soap to use). This social information is even

more important when the appropriate behavior is unclear or when situations are ambiguous or uncertain, as extant research has shown that social norms affect behavior to a greater degree in such settings (Festinger, 1954; Sherif, 1936). In effect, descriptive social norms can function as heuristics for decision making, providing a guide for appropriate or wise behavior in a wide range of situations.

Together, both by conveying what is appropriate and likely to be effective, descriptive social norms produce powerful effects on judgments and decisions (Cialdini, 2003; Cialdini, Reno, & Kallgren, 1990). A large body of empirical evidence has shown that descriptive social norms serve as salient reference points for behavior in many contexts, ranging from energy consumption to job acceptance decisions (Coffman et al., 2014; Goldstein et al., 2008; Nolan et al., 2008). We propose that descriptive social norms should influence decisions made about group diversity just as they influence decisions in other contexts. Past research on scrutiny and impression management illuminates why those responsible for decisions influencing group diversity may feel pressure to follow descriptive social norms.

How Scrutiny of Group Diversity May Drive Conformity to Descriptive Social Norms

Scrutiny refers to obtrusive and critical attention paid to particular behaviors (Sutton & Galunic, 1996), and scrutiny can come from a variety of sources. For example, the media is one common source of scrutiny capable of influencing an organization's reputation and value and shaping others' perceptions of its legitimacy. Naturally,

organizations compete to receive positive and avoid negative media exposure (Fombrun, 1996; Fombrun & Shanley, 1990; Pollock & Rindova, 2003). Scrutiny can also come from other sources such as shareholders (e.g., institutional investors placing pressure on firms to engage in socially responsible behaviors) and policy makers (e.g., through regulations and the imposition of rewards or penalties for certain behaviors; Aguilera, Rupp, Williams, & Ganapathi, 2007; Campbell, 2007). The public also often directly scrutinizes organizations, mobilizing in ways that may draw wanted or unwanted attention to particular behaviors (e.g., through social movement boycotts; McDonnell & King, 2013).

In general, groups and organizations have strong incentives to avoid negative scrutiny. Negative scrutiny can be detrimental for reputation and legitimacy (Desai, 2011), so in order to avoid negative scrutiny, groups frequently attempt to manage impressions around scrutinized behaviors (Bolino, Kacmar, Turnley, & Gilstrap, 2008; Elsbach, Sutton, & Principe, 1998). Impression management describes attempts by groups or organizations to positively shape how they are perceived (Elsbach & Sutton, 1992), and it may occur even in anticipation of the possibility of negative events. For example, Elsbach et al. (1998) have documented how hospitals use anticipatory impression management tactics in order to prevent potential negative scrutiny.

In recent years, scrutiny has increased surrounding the diversity of groups. For example, the media has scrutinized companies for insufficient gender diversity on their boards of directors (Merchant, 2013; Miller, 2013); presidents for insufficient race and

gender diversity in their cabinets (Lee, 2017) and their U.S. Supreme Court nominees (Totenberg, 2016); and the Academy of Motion Picture Arts and Sciences for insufficient racial diversity among their Oscar nominees (Buckley, 2016; Ryan, 2016). Importantly, scrutiny is often applied selectively: rather than simultaneously emphasizing racial, gender, and socio-economic diversity, for instance, scrutiny often focuses more narrowly on a single dimension of diversity. For example, while groups such as corporate boards have faced considerable negative scrutiny for a lack of gender diversity, there has been far less attention to their lack of racial diversity.

Scrutiny surrounding diversity naturally motivates impression management concerns. An important question, then, is how decision makers who shape the composition of high profile groups within organizations may seek to manage diversity in order to avoid negative scrutiny. We propose that past research on descriptive social norms provides key insights. If groups or organizations are motivated to avoid negative scrutiny, then following the descriptive social norm for diversity essentially ensures that they will not be singled out for inadequate diversity. Further, because it is often unclear what an "objective" benchmark for strong performance should be in the context of decisions around diversity (Bell & Hartmann, 2007; Shemla, Meyer, Greer, & Jehn, 2014; Unzueta et al., 2012), descriptive social norms should be particularly informative in guiding behavior around diversity. Thus, groups and organizations (and the decision makers responsible for their composition) may treat the descriptive social norm for diversity as a goal for impression management reasons.

The Implications of Descriptive Social Norms as Diversity Goals

Past research on goal setting offers insight into what will happen when those who shape group composition share the same explicit or implicit goal. Goals serve as reference points, causing individuals to expend considerable effort in the hopes of achieving an unmet goal and then to relax their efforts after achieving it (Heath et al., 1999; Locke & Latham, 2002). This has been shown to lead to performance clustering around salient goal thresholds in numerous contexts. For instance, professional baseball players finish seasons disproportionately often with a batting average just above .300 (a salient threshold widely believed to separate good hitters from great ones; Moskowitz & Wertheim, 2011; Pope & Simonsohn, 2011), and marathon runners finish races disproportionately often in the minute right before salient, round number thresholds (e.g., the minute just under three hours; Allen et al., 2016). We therefore expect to observe an excess mass or clustering of groups at (or just above) the descriptive social norm for diversity.[1]

Hypothesis 1a. Groups' diversity levels will cluster at (or just above) the descriptive social norm set by peers for diversity.

While Hypothesis 1a pertains to group composition, group composition is the result of decisions regarding which members to add to a group. If reaching the descriptive

[1] Because descriptive social norms are averages, they are rarely whole numbers (e.g., the average number of women per board was 1.36 women in the S&P 1500 in 2013). Since groups cannot have fractional numbers of women or racial minorities, we expect clustering at "or just above" the descriptive social norm (i.e. at the smallest whole number above the descriptive social norm).

social norm for diversity is a goal of those who shape group compositions, then efforts to increase group diversity (in the form of adding underrepresented group members) should decline precipitously once the descriptive social norm for diversity is achieved. Empirically, this relaxing of effort after reaching a goal threshold has been observed in several contexts. In the context of baseball, as just mentioned, batters and their teams reduce their at bat appearances near the end of the season once they have exceeded the salient .300 batting average threshold that separates good hitters from great ones (Pope & Simonsohn, 2011). In the context of SAT scores, students are disproportionately less likely to retake the SAT once they surpass a salient threshold such as a score of 1000 (the average score set by the College Board and a salient round number; Pope & Simonsohn, 2011). In our context of diversity and group composition decisions, we predict that groups are less likely to increase their diversity once they have already reached the descriptive social norm for diversity established by peers.

Hypothesis 1b. Groups (and the individuals who shape their composition) will add new members from underrepresented populations at a lower rate once they have surpassed the pertinent descriptive social norm for diversity.

Importantly, we only expect descriptive social norms to serve as goals when it comes to scrutinized dimensions of diversity. Without any scrutiny on a given dimension of diversity, there should be no impression management motives and thus no desire to follow the descriptive social norm. For example, we would expect to find support for Hypotheses 1a and 1b when it comes to gender diversity in settings where inadequate

gender diversity has been scrutinized (e.g., on corporate boards) but not in settings where gender diversity has not been scrutinized. Thus, we propose that scrutiny (or the threat of negative scrutiny) is required in order to produce our hypothesized clustering and threshold effects.

Hypothesis 2. Scrutiny moderates the effects of descriptive social norms on group diversity decisions. Specifically, descriptive social norms will only influence group diversity decisions and outcomes when scrutiny is present along a given diversity dimension.

The Moderating Role of Visibility

If groups and organizations manage impressions around diversity to avoid negative scrutiny, this tendency should be more pronounced among more visible groups and organizations. We follow past research and use the term "visibility" to describe how much attention individuals, groups, or organizations typically receive (Chiu & Sharfman, 2011), regardless of why they are receiving this attention (as opposed to our use of the term "scrutiny", which refers to attention paid to a particular behavior such as a group's gender diversity). When firms are more visible (e.g., because they operate in more visible industries or because they have higher overall media exposure), they face greater external pressures to engage in legitimacy-seeking behaviors (Gardberg & Fombrun, 2006) and are also more likely to engage in legitimacy-enhancing behaviors like corporate social performance initiatives (Chiu & Sharfman, 2011). For example, firms respond more to boycotts when they receive more media attention (King, 2008), and firms engage in more

prosocial activities when boycotts are more threatening because of increased media attention (McDonnell & King, 2013). Past research has shown that conforming to descriptive social norms (i.e. mimicking the behavior of peer firms) is one way to enhance legitimacy (DiMaggio & Powell, 1983), suggesting that descriptive social norms should influence the diversity of groups along scrutinized diversity dimensions to a greater degree when those groups are more visible. Further, the actions of more visible firms receive more attention, which can magnify the negative consequences of failing to conform to social norms.

Past research on individual judgment and decision making makes similar predictions regarding the effects of visibility on conformity to descriptive social norms. Social norms influence behavior to a greater degree when individuals and their behaviors are more visible (Cialdini & Trost, 1998). In particular, individuals tend to look to social norms to guide their behavior most frequently when the behavior in question is public or observable (Cialdini, Kallgren, & Reno, 1991; Cialdini et al., 1990; Kallgren, Reno, & Cialdini, 2000; Shaffer, 1983). For example, studies have found that monitoring employees can improve conformity to ethical norms in the context of employee theft (Pierce, Snow, & McAfee, 2015), monitoring can improve conformity to hand hygiene norms in hospitals (Staats, Dai, Hofmann, & Milkman, 2016), and being in a public setting (as opposed to a private setting) can make women more likely to conform to gender norms regarding assertiveness (Swim & Hyers, 1999). On an individual level, we would thus expect more conformity to descriptive social norms when outcomes are more

visible. Thus, research and theorizing on both individuals and firms suggests more visible groups should be more likely to conform to social norms around diversity along scrutinized diversity dimensions.

Hypothesis 3. Visibility moderates the effects of descriptive social norms on group diversity decisions along scrutinized diversity dimensions. Specifically, more visible groups will be more likely to follow the descriptive social norm for diversity along scrutinized diversity dimensions than less visible groups.

OVERVIEW OF STUDIES

The remainder of this paper proceeds as follows. We begin by examining our hypotheses in the field, exploring whether they make accurate predictions about the composition and evolution of U.S. corporate boards. In Study 1A, we present analyses of S&P 1500 board composition data from 2013 that test for excess clustering of corporate boards at the descriptive social norm for gender diversity (Hypothesis 1a). We also examine whether this pattern is more extreme among more visible companies (Hypothesis 3). In Study 1B, we present analyses of board member additions to determine whether boards are discontinuously less likely to add female directors once they have reached the descriptive social norm for gender diversity (Hypothesis 1b). In Study 1C, we run an online experiment to test for evidence of the same pattern of discontinuities in board member selection found in the field in Study 1B in a stylized hypothetical decision environment where we can randomize the number of women on a board and control for the availability of qualified candidates (Hypothesis 1b). In Studies

2A and 2B, we seek evidence that scrutiny, descriptive social norms about diversity, and goal thresholds influence the gender of group members selected for empty positions, and we experimentally manipulate social norms and scrutiny to test Hypotheses 1b and 2. Finally, in Study 3, we examine how social norms and group visibility affect the race of group members selected for empty positions, and we do this by experimentally manipulating social norms and visibility to test Hypotheses 1b and 3. Together, these studies help establish the external validity, internal validity, and generalizability of our theories.

STUDY 1: CORPORATE BOARDS

We first test our theories in the context of U.S. corporate boards. This is an important organizational setting that is economically significant, as boards control trillions of dollars. It is also highly policy relevant, as in recent years, numerous countries have passed laws about the gender composition of the corporate boards of public companies (Bainbridge & Henderson, 2014; Forbes & Milliken, 1999; Smale & Miller, 2015).

Study 1A: Clustering of U.S. Corporate Board Compositions at the Social Norm

In Study 1A, we analyzed the most recent available S&P 1500 corporate board composition data (from 2013) to test whether descriptive social norms influence board composition. Given the importance of scrutiny to our theoretical model (see Hypothesis 2), we first sought to establish which dimensions of corporate board diversity faced scrutiny at the time of data collection. An analysis of news articles from 2013 in the news

database Lexis Nexis revealed that of 98 newspaper articles that mentioned "board diversity", 97% mentioned gender diversity, while 18% mentioned racial or ethnic diversity (the second most frequently mentioned social category). In addition, several countries in Europe have recently passed laws mandating minimum levels of gender diversity on the boards of public companies under their jurisdiction (Smale & Miller, 2015), but no such laws have been passed about other types of diversity. Given that the majority of attention regarding diversity on corporate boards focuses on gender diversity, in this study, we therefore test for (and only expect to observe) social norm effects pertaining to the gender diversity of U.S. corporate boards.

On S&P 1500 corporate boards, the average number of women was 1.36 in 2013, and this descriptive social norm received significant media coverage, with all newspaper articles in the Lexis Nexis database about board gender diversity in 2013 focusing on the average number or percentage of women on boards. We therefore expect to observe an excess of boards with exactly two women, as boards with two women just exceed the peer norm for gender diversity (Hypothesis 1a). We also predict that this excess of exactly two women per board will be more prevalent among more visible companies– those that receive more overall media attention (Hypothesis 3).

Methods

Data. Our dataset was compiled by Institutional Shareholder Services (ISS). The ISS Directors Data we analyzed contains detailed information about the boards of directors for 1,514 companies that represent the S&P Composite 1500, which is

composed of three indices: the S&P 500, the S&P MidCap 400, and the S&P SmallCap 600. The S&P 1500 represents roughly 90% of the total U.S. stock market capitalization, and we also focus on the far more visible subset of companies in the S&P 500,[2] which represents roughly 90% of the total market capitalization of the S&P 1500 and 80% of the total market capitalization of the U.S. stock market (S&P Dow Jones Indices, 2015).

The ISS dataset we analyze includes information on the individual members of the boards of directors for each of the 1,514 companies in the S&P Composite 1500, including each director's name, gender, and ethnicity.[3] The dataset is updated annually, and for our primary analysis, we relied on the 2013 data, as this was the most recent data available to us as of June 5, 2015 when we first accessed the ISS database.

Additional data were collected on each company's media mentions (from Lexis Nexis), industry (from NASDAQ), year of IPO (from Bloomberg and company websites), market capitalization (from the Center for Research in Security Prices and Google Finance), and percent institutional ownership (from Bloomberg), and these data were used to perform robustness checks and investigate the moderating effect of visibility.

---

[2] A Google search for the term "S&P 500" returns 400 times as many results as a Google search for the term "S&P 1500", and a Google Scholar search for the term "S&P 500" returns 20 times as many articles as a search for the term "S&P 1500".

[3] ISS data on director gender was complete, but in 31 instances, director ethnicity was missing or blank. We manually searched Google and company websites to fill in these missing data.

Analysis Strategy. To test Hypothesis 1a, we relied on a comparison of the *actual* distribution of male and female directors on corporate boards with the distribution we would *expect* if those directors were assigned to boards in a gender-neutral manner. We determined the expected distribution using a Monte Carlo simulation method (Rubinstein & Kroese, 2011). Specifically, we took existing 2013 S&P 1500 and S&P 500 data on directors and board seats from the ISS Directors dataset and then randomly reassigned directors to different boards, generating 10,000 simulated distributions of directors to boards. Because we *randomly* reassigned actual directors to boards in each of our simulations, these simulations produced the board composition distribution we would expect to see if gender played no role in board member selection. In other words, given the available pool of board seats and directors, our simulations told us how many women we should expect to see on each board if boards ignored gender when selecting board members.

We reassign *existing* directors in our simulations to provide a conservative test of whether there exist anomalous sorting patterns of female directors to boards.[4] In each

[4] One common explanation for the limited number of women on corporate boards is that there are not enough qualified women to serve on boards. We thus assume the universe of people qualified to serve on boards consists only of those who actually sit on boards, so our simulations gauge whether we find anomalous sorting even if we assume no more qualified women exist to serve on boards. This extremely conservative assumption is certainly incorrect, but given that the universe of qualified women must be larger than the set who already serve on boards, finding evidence of clustering at the social norm under our assumptions would be even *more* remarkable (since relaxing this assumption would make it easier for the observed gender distribution to deviate from our simulated expected distribution).

simulation, we took as given the number of boards, the size of each board, and the number of board seats each director held based on the statistics we observed in the 2013 ISS Directors Dataset. For example, if company Alpha had nine board members in the ISS Directors Dataset, then in each simulation, company Alpha was assigned nine distinct board members. Similarly, if director Zed held two different board seats in the ISS Directors Dataset, then director Zed ended each simulation holding seats on two different corporate boards.

Running this simulation 10,000 times produced random assignments of all directors to all boards that reflected the same number of directors, number of boards, and the same board sizes we observed in the ISS Directors Dataset. For each simulation result, we considered how many company boards were assigned zero female directors, one female director, two female directors, etc. We then calculated the mean of these values across all 10,000 simulations. These means told us how many companies we would expect, on average, to observe with exactly zero, one, two, and so on female directors if available board seats in the ISS dataset were randomly assigned to available directors. Our simulations also told us how rare a given assortment was, giving us bounds in the form of confidence intervals around each mean to indicate the likelihood under random assignment that we would observe a certain fraction of boards containing a specific number of women (e.g., in what fraction of 10,000 simulations had we obtained such a result).

Although this simulation strategy has been used and validated in a number of empirical papers (e.g., Dezső, Ross, & Uribe, 2016; Gino & Pierce, 2010), we also conducted placebo simulations with a characteristic other than gender to ensure that any observed deviations from our simulations on gender were not an artifact of our simulation method (see *Robustness Checks*).

Results

Summary Statistics. For companies in our dataset, the modal number of directors on a board was nine, the median number was nine, and 95% of companies had between 6 and 14 directors. Because we were interested in understanding the distribution of the absolute number of women on each board, boards with outlier numbers of seats could have exerted undue influence on our analyses. For our primary analyses, we therefore trimmed our dataset to include only companies with a total number of directors in the middle 95% of the distribution, excluding companies with outlier numbers of directors (i.e. fewer than six or more than 14) and leaving us with 1,441 companies to analyze. However, the results of our analyses remain meaningfully unchanged in terms of magnitude and statistical significance if we repeat them without trimming these outliers (see *Online Supplement*).

The 1,441 companies in our trimmed data set included 13,440 distinct board seats and 11,185 distinct directors, as some directors held board seats on multiple company boards. In our trimmed dataset, 84% of directors held exactly one board seat; 13% held two board seats; 3% held three board seats; and less than 1% held four or five board

seats. Of the 11,185 unique directors represented in our trimmed dataset, 14% (1,558) were female, and women held 15% (1,963) of the available board seats (see Table 1). Ninety-one percent (N = 10,150) of directors were Caucasian, 3.7% (N = 417) were Black, 3.0% (N = 335) were Asian, 1.7% (N = 192) were Hispanic, and 0.8% (N = 91) were classified as belonging to a different ethnic group (see Table 1). The average age of the directors in our trimmed dataset was 62.9 years with a standard deviation of 8.9 years. Fifty-eight (4.0%) of the companies had female CEOs. See Table 2 for a correlation matrix describing our data.

Do Boards Cluster Around the Descriptive Social Norm for Gender Diversity?

Hypothesis 1a suggests we should find an excess of boards with exactly two women (since the relevant descriptive social norm was that an average board in the S&P 1500 included 1.36 women in 2013 and an average board in the S&P 500 included 1.89 women in 2013). Based on simulations of the S&P 1500, there were 8% fewer companies with no women than would be expected ($p < 0.02$), and consistent with Hypothesis 1a, there were 12% more boards with exactly two women than would be expected ($p < 0.01$). Boards including other frequencies of women were in line with expectations (see Figure 1, Panel A). Similarly, for the S&P 500 and consistent with Hypothesis 1a, there were 45% more companies with exactly two female board members than would be expected ($p < 0.001$). There were also 45% fewer companies with no female board members than we would expect ($p < 0.001$), and boards including other frequencies of women again arose at the rate expected (see Figure 1, Panel B). Thus, Hypothesis 1a is supported, and in light of

the far higher visibility of S&P 500 companies than other companies in the S&P 1500, these patterns provide suggestive evidence in support of Hypothesis 3.

To provide further support for Hypothesis 1a, we analyzed additional historical data on corporate board composition to assess whether historical descriptive social norms also determined where clustering occurred. In years when the average number of women per board (i.e. the descriptive social norm) was below one woman, our theorizing predicts an overrepresentation of boards with exactly one woman (i.e. "tokenism" or a group including exactly one woman (Kanter, 1977)); in years when the average number of women per board was between one and two women (e.g., 1.36 women per board in 2013), our theorizing predicts an overrepresentation of boards with exactly two women. We name the phenomenon whereby a group includes exactly two women "twokenism", which is a portmanteau of the number "two" and the term "tokenism" originally used by Kanter (1977). We repeated our simulations using twelve years of historical data to see if the descriptive social norm did in fact predict where an excess of boards arose in each distribution.

We gathered additional data on the composition of S&P 1500 boards from 2002 to 2012 from the RiskMetrics Directors Legacy dataset (for the years 2002 to 2006)[5] and

[5] Data captured prior to 2002 in the RiskMetrics Directors Legacy dataset appear to have substantial variation in data quality and reliability. For example, although the dataset is meant to include information about S&P 1500 companies, and there are roughly 1500 companies in the S&P 1500, the 2001 dataset included information about 1797 companies supposedly in the S&P 1500, suggesting it was unreliable. This is why we

the ISS (RiskMetrics) Directors dataset (for the years 2007 to 2012) on August 22, 2016. For each year from 2002 to 2012, we repeated our simulation strategy to calculate how many boards would be expected to include exactly one or exactly two female directors and then compared these simulation-based expectations to the number of boards we actually observed with exactly one or exactly two female directors.

As illustrated in Figure 2, we found a statistically significant overrepresentation of boards with exactly one woman when the descriptive social norm was below one woman per board and statistically significant overrepresentation of boards with exactly two women when the descriptive social norm rose above one woman per board. In 2002 and 2003, the descriptive social norm for gender diversity–or the average number of women per board–was less than one woman, and we see statistically significant tokenism in these two years, but we do not find statistically significant twokenism in these years. From 2005 to 2013, the descriptive social norm for gender diversity exceeded one woman, and we see statistically significant twokenism in these years, but we do not find statistically significant tokenism in these years, however. In 2004, the first year that the descriptive social norm for gender diversity exceeded one woman in the S&P 1500, we still observe statistically significant tokenism and do not yet find statistically significant twokenism.

When we ran an ordinary least squares regression with robust standard errors clustered at the firm level to predict the extent of tokenism (or the overrepresentation of

---

began our analyses with data from 2002. ISS Director data is only available going back to 2007.

boards including one woman) or twokenism (or the overrepresentation of boards including two women) in each year as a function of whether the descriptive social norm for gender diversity exceeded one woman in that year, we found that the descriptive social norm exceeding one woman was a significant *negative* predictor of tokenism (β = -0.11; $p < 0.001$) and a significant *positive* predictor of twokenism ($\beta = 0.12$; $p = 0.002$). This provides further support for Hypothesis 1a and our theorizing that descriptive social norms help determine salient thresholds for diversity.

Are More Visible Companies More Likely to Exhibit Twokenism? To test Hypothesis 3 in this context, we examined whether companies that receive more media attention were more likely to include exactly two women on their boards. We used media attention as a proxy for visibility to align with past research on organizational visibility (Brammer & Millington, 2006; Chiu & Sharfman, 2011; King, 2008; McDonnell & King, 2013). We searched Lexis Nexis for all media mentions (including newspapers, web-based publications, magazines, etc.) of each of the companies in the S&P 1500 in 2012 (mean media mentions of a company = 307; S.D. = 441). We gathered 2012 data on media attention so we could examine whether past media attention predicted future (2013) twokenism. We then analyzed whether media attention in 2012 predicted whether companies would include exactly two women on their boards in 2013.

We ordered the companies in our dataset by the number of media mentions each company received in 2012 and created deciles (i.e. ten bins of 144 companies each) based on this ordering. Thus, the first decile contained the companies most frequently

mentioned in the media in 2012, while the last decile contained the companies least frequently mentioned in the media in 2012. After segmenting the companies in our dataset by the amount of media attention they were subjected to in 2012, we repeated our basic simulation strategy but limited each simulation to include only the companies in a given decile. This allowed us to determine how many companies we would expect to see with exactly two women on their boards in 2013 in each of the deciles. We ran 1,000 simulations for each decile, generating a new expected number of companies with exactly two female directors each time. Thus, for each decile, we generated an *expected* number of companies with exactly two women on their boards based on our simulations, and we could compare this with the *actual* number of companies including exactly two women on their boards in our 2013 board data.

The results of our simulations for the different media attention deciles are depicted in Figure 3. To test the hypothesis that the likelihood of having exactly two women on a company's board increases for more visible companies, we ran an ordinary least squared (OLS) regression with robust standard errors. We used the logarithm of the average number of media mentions in a given decile to predict the absolute difference between the observed and expected number of companies with exactly two women on their boards in each decile. The logarithm of media mentions of the decile was a significant predictor of the absolute difference between observed and expected boards with exactly two female directors ($\beta = 6.12$, $p = 0.014$). The positive coefficient of log

media mentions indicates that deciles containing more visible companies were more likely to display twokenism, supporting Hypothesis 3.[6]

Robustness Checks. To further validate our simulation strategy and ensure our results were not an artifact of the way we constructed an expected distribution of the number of boards including varying numbers of female directors, we conducted a series of placebo simulations (Gino & Pierce, 2010). Specifically, in these placebo simulations, we produced expected distributions of the number of boards that would include varying numbers of directors with another characteristic (i.e., not gender) that should not show goal-related clustering effects because of a lack of scrutiny on that characteristic (e.g., board members whose ages ended with an arbitrary number). We found no significant differences between the expected numbers of boards and the actual numbers of boards in any of our placebo simulations, suggesting that the large deviations we see in our simulations studying gender were not an artifact of the way we constructed our baseline expectations or null distributions (see the *Online Supplement* for complete details about our placebo simulations).

In addition to conducting placebo simulations to ensure the robustness of our simulation methodology, we conducted numerous additional robustness checks to ensure

---

[6] See the *Online Supplement* for additional specifications of this regression to test the robustness of this finding and for a table reporting detailed regression results. We used as predictors either the logarithm of the average number of media mentions or the decile rank, and we used as outcomes either the absolute overrepresentation of boards with exactly two women or the percent overrepresentation of boards with exactly two women. All yielded findings that were statistically significant and meaningfully unchanged.

our results were not driven by outliers or by a small subset of boards by repeating our baseline simulations with different cuts of our data. First, we checked that our findings were robust to board size. To do this, we used our standard simulation strategy but limited the data to boards of size 6 or fewer, 7, 8, 9, 10, 11, 12, or 13 or more. The underrepresentation of companies with no women on their boards and the overrepresentation of companies with exactly two women on their boards is robust across all board sizes tested (see Table 3), although the clustering at the social norm of two is largely driven by companies with larger boards, and future research exploring the reasons for this could yield interesting insights.

Our results are also robust across industries, and they hold regardless of the gender of a company's CEO (see *Online Supplement*). We also examined whether the length of time the company has been public affects the likelihood that the company has exactly two women on its board. In general, our results appear to be robust regardless of when a company went public (see *Online Supplement*). Finally, when we examine how our results relate to the market capitalization of a company, we find that twokenism is more prevalent among companies with higher market capitalization (see *Online Supplement*), which are also the most frequently mentioned by the media (the correlation between the logarithm of a company's market capitalization and the logarithm of its number of media mentions in 2013 = 0.59; $p < 0.001$).

Study 1B: Threshold Effects in Board Member Selection at the Social Norm

In Study 1B, we analyzed the gender of new board members added to company boards over time for evidence consistent with our theories. We predicted that boards would be discontinuously less likely to add additional women once they had met the relevant descriptive social norm for gender diversity (Hypothesis 1b). Given that the descriptive social norm for gender diversity in the S&P 1500 first surpassed one woman in 2004, we examine all board member additions from 2004 to 2013 to test whether boards during this time period were discontinuously less likely to add additional female directors once they already included two women on their boards.

Method

Data. For these analyses, we use a subset of the data described in Study 1a. Specifically, we use the ISS Directors dataset describing board composition from 2007 to 2013 and the RiskMetrics Directors Legacy dataset describing board composition from 2004 to 2006 to examine the 9,989 board member additions in the S&P 1500 from 2004 to 2013.

Analysis Strategy. Using data on all board member additions from 2004 to 2013, we estimated an ordinary least squares regression with robust standard errors to predict whether each newly added board member was female.[7] We included as predictors both the number of women currently on a board (to control for the possibility that boards have

[7] We rely on a linear model because it yields more interpretable coefficients than a logit specification, and this method also allows us to correct for heteroskedasticity in the standard errors (Angrist & Pischke, 2008; see Brands & Fernandez-Mateo, 2016 for a similar procedure). However, logistic regressions yield similar results and are reported in the *Online Supplement*.

either increasing or decreasing marginal value for female directors) as well as an indicator for whether the board included at least two women (our primary predictor of a discontinuity in a groups' desire to add more women after exceeding the social norm for gender diversity), and we clustered standard errors by firm. We report these regressions with and without fixed effects for board size, fixed effects for industry, fixed effects for stock market index, and a continuous control for a company's market capitalization.

Results

Summary Statistics. Of the 9,989 board additions from 2004 to 2013, 16.5% (1,649) were additions of female directors. The 9,989 board member additions from 2004 to 2013 represent 8,328 distinct directors (i.e. some directors were added to multiple boards during this time period), and 16.2% (1,347) of the distinct directors were female. On average, boards in this dataset added 5.25 directors during this nine-year span.

Do Boards Add Fewer Women Once They Have Reached the Descriptive Social Norm? As shown in Table 4, Model 1, for the S&P 1500, the coefficient on our primary predictor of whether a board added a female director–an indicator for whether the board already included at least two women–was negative and significant in our primary regression specification (β = -0.039; p = 0.017). As shown in Table 4, Model 3, for the S&P 500 (roughly the 500 most visible and valuable companies in the S&P 1500), the coefficient on the indicator variable was negative and even more highly significant (β = -0.092; $p < 0.001$). This suggests that companies are less likely to add additional women to their boards once their boards have met the social norm for gender diversity by

including two women, providing support for Hypothesis 1b. The larger effect size in the (highly visible) S&P 500 also provides some suggestive support for Hypothesis 3. Adding in fixed effects for board size, fixed effects for industry, fixed effects for stock market index, and a continuous control for market capitalization (see Table 4, Models 2 and 4), we still find that our predictor of a discontinuity is significant in the S&P 1500 (β = -0.034; p = 0.037) and in the S&P 500 (β = -0.090; $p < 0.001$).

Do More Visible Companies Show Larger Discontinuities at the Descriptive Social Norm? To test Hypothesis 3 in Study 1B, we examined whether there was an interaction between media attention and our primary predictor of whether a board added a female director–an indicator for whether the board already included at least two women. We again searched Lexis Nexis for all media mentions of each of the companies in the S&P 1500, and we gathered additional data to look at media mentions for each year starting in 2004 to see if media attention in year $t - 1$ predicted whether a newly added board member in year t was female. For our analyses, we used the centered logarithm of media mentions rather than the raw number of media mentions because the distribution of media mentions is highly skewed (skewness = 2.57; skewness test for normality $p <$ 0.0001; kurtosis = 10.37; kurtosis test for normality $p < 0.0001$).

Our results are depicted in Table 5. As predicted, we find a significant negative interaction between the centered logarithm of the number of media mentions in year $t - 1$ and having two or more women on a board in predicting whether a newly added board member in year t was female (β = -0.021; p = 0.042; Model 1). Adding in fixed effects

for board size, fixed effects for industry, fixed effects for stock market index, and a continuous control for market capitalization, we still find a significant negative interaction between the centered logarithm of media mentions and having two or more women (β = -0.021; p = 0.041; Model 2). These results suggest that more visible companies show larger discontinuities in board member additions at the descriptive social norm of two women per board.

Study 1C: Online Experiment Replicating Threshold Effects

In Study 1C, we sought to replicate our findings regarding threshold effects from Study 1B in an online experiment that allowed us to randomly assign the number of women in a group and control for the availability of qualified candidates. Specifically, we investigated whether individuals in a controlled setting are less likely to add women to a corporate board when the board has met or surpassed the social norm for gender diversity by including two or more women.

Method

Participants. Four hundred and seventy-nine U.S. participants were recruited through Amazon's Mechanical Turk to participate in a short online research study (55% male; 77% Caucasian). These participants were paid $0.25 for completing a survey they were told would take approximately 5 minutes of their time. Sample size was determined a priori, data analysis was conducted only once all data were collected, and we do not exclude any data.

Procedures. In a pilot study (see the *Online Supplement* for details), we first established that our study population was indeed aware that two is the average number of women on U.S. corporate boards (i.e. two women is the descriptive social norm for gender diversity).

After establishing an awareness of descriptive social norms in a pilot, we ran our primary study. In this study, participants were asked to imagine they had been tasked with helping a company select a new member for its board of directors. They were then exposed to a list of ten names and told the current board consisted of the individuals on that list. Participants were randomly assigned to one of four experimental conditions where zero, one, two, or three of the names of board members were female.

Study participants were next presented with three hypothetical candidates for an opening on the board in question and asked to choose one to add to the board. The candidates were all described as qualified, but one was a CEO, one was a current board member at another company, and one was a consultant with expertise in the industry. We randomly varied which candidate had a female name (Jill Davis) and which candidates had male names (Matthew Anderson and Todd Miller), and we randomly varied which name was associated with each qualification.[8] We presented three candidates for the available board seat rather than one male and one female to reduce suspicion that our

---

[8] Participants were most likely to choose the candidate who was a CEO ($p < 0.001$), regardless of gender. However, because we randomly assigned qualifications to the candidates, we do not need to control for candidate qualification in order for our tests to provide unbiased estimates of the causal effects of our manipulations. In addition, we did not find any significant interactions between gender and candidate qualifications.

study was about gender following Castilla and Benard (2010). Our dependent variable of interest was what fraction of participants in each condition would choose a female candidate.

Finally, participants completed demographic questions and a manipulation check question, which asked them to recall how many men and how many women were present on the corporate board they had seen at the beginning of the survey. Study materials and a correlation matrix of all variables collected in this study are available in the *Online Supplement*.

Results

First, our manipulation check confirmed our manipulation was successful: participants recalled significantly more women on boards that included three women than two ($p < 0.001$), two women than one ($p < 0.001$), and one woman than zero women ($p = 0.015$).

Second, a Chi-square test of independence showed a marginally significant relationship between the number of women on the board and whether the participant chose the female candidate ($\chi^2(3, N = 479) = 7.51$, $p = 0.057$). Consistent with Hypothesis 1b and replicating our results from Study 1B, participants were significantly less likely to choose the female candidate and increase the gender diversity of the board once the board included at least two women. Participants shown a corporate board with exactly two female members were significantly less likely to choose the female candidate for the open seat ($M = 36.0\%$, $SD = 0.482$) than were participants who were shown a

corporate board with one female member (M = 50.4%, SD = 0.502; $t(239) = 2.27$, p = 0.024; see Figure 4).[9] We then ran an ordinary least squares regression (OLS) with robust standard errors to predict the likelihood a participant chose the female candidate, replicating our empirical analyses of board member additions from S&P 1500 and S&P 500 data from Study 1B. We again included the number of women currently on the board as a control variable in addition to an indicator variable for whether the board included at least two women as a predictor of a discontinuity. The coefficient on the indicator variable was negative and marginally significant (β = -0.187, p = 0.062; see Table 6), suggesting that participants in our experiment also were discontinuously less likely to increase the gender diversity of the board once the board had at least two women and providing additional support for Hypothesis 1b.

Discussion

Study 1A shows that U.S. corporate boards are disproportionately likely to include exactly the number of women needed to minimally exceed the descriptive social norm for female representation in peer groups. This evidence is consistent with Hypothesis 1a, which proposes that the composition of groups facing scrutiny along a diversity dimension will cluster around the descriptive social norm for that type of diversity. Further, historical analyses show that descriptive social norms predicted the shift from tokenism (an overabundance of boards with exactly one female director) to

---

[9] We find a main effect of participant gender such that female participants are significantly more likely to select the female candidate (p = 0.019), but we find no significant interaction between participant gender and decisions.

twokenism (an overabundance of boards with exactly two female directors), providing additional support for Hypothesis 1a that the clustering we detect is driven by the descriptive social norm for gender diversity.

Study 1B provides support for Hypothesis 1b, which states that groups facing scrutiny along a diversity dimension will be less likely to add members of the relevant underrepresented group once they have reached the descriptive social norm for diversity. We find that U.S. corporate boards are discontinuously less likely to add additional women once they have reached the descriptive social norm for diversity by including two female directors. In Study 1C, we replicate this finding in a stylized experiment where we randomly assign the number of women to a hypothetical corporate board and control for the availability of qualified candidates. While Study 1C lacks the realism of Studies 1A and 1B, it confirms our hypothesis in an environment where we can randomly assign board composition, providing convergent evidence that there exists a causal relationship between board composition and the gender of new board members.

Consistent with Hypothesis 2, which predicts scrutiny is a necessary condition for social norms to influence diversity, we do not see evidence of clustering at the social norm when we look at board members' race or ethnicity in supplemental analyses.[10] There is far less scrutiny of corporate boards' racial diversity compared with the scrutiny boards face regarding gender diversity (i.e., only 18% of news articles about board

---

[10]A more detailed discussion of simulation analyses regarding director race and ethnicity can be found in the *Online Supplement*.

diversity in 2013 discussed racial diversity while 97% discussed gender diversity, and no laws have been passed establishing racial quotas on corporate boards in any country), so corporate boards may have fewer impression management motives regarding the recruitment of racial or ethnic minorities compared to women.

Finally, consistent with Hypothesis 3, we find evidence that more visible companies (as measured by media coverage in the previous year) are more likely to include exactly two women on their boards, consistent with our theory that the clustering we detect at the social norm is driven in part by impression management concerns. In Study 1B, we also find that more visible companies show larger discontinuities at the descriptive social norm of two women per board when adding additional female board members.

Past research suggests these findings are worrisome from a policy perspective. Research on the benefits of gender diversity on corporate boards suggests that at least three female directors are needed before boards experience tangible benefits from gender diversity (Konrad, Kramer, & Erkut, 2008; Torchia, Calabrò, & Huse, 2011). By stopping at two women, boards may be missing out on key benefits that can ensue from greater gender diversity (cf. Adams & Ferreira, 2009; Matsa & Miller, 2013). Further, our results suggest that the push for gender parity on boards may not generate results for a long time. In Study 1A, we depict the evolution of descriptive social norms regarding gender diversity on corporate boards over a twelve-year span, and these results suggest that descriptive social norms change quite slowly over time.

In spite of the compelling evidence provided by our empirical analyses of archival board composition data supporting our theorizing and hypotheses, Studies 1A and 1B are ultimately only correlational studies and thus have limitations. We cannot completely rule out concerns about reverse causality or other confounds such as firm performance. In addition, because we do not observe board member selection decisions directly, we can only explore the mechanisms responsible for the effects we have documented indirectly. There are many factors at play that affect who is added to corporate boards (e.g., legal constraints can prevent people from serving on multiple boards; bias and stereotyping may affect board member selection), and we focus only on the roles played by descriptive social norms, scrutiny, and visibility. We also unfortunately cannot disentangle the specific motives of individual companies.

In order to provide more confidence in our results, in Study 1C, we replicated threshold effects at the descriptive social norm in an experimental setting where we could randomize the number of women in a group. This gives us greater confidence that the results found in Study 1B are not driven by endogeneity or the fact that there are not enough qualified women for director positions. However, we acknowledge that Study 1C is a stylized experiment that does not accurately represent corporate board decision-making processes. First, our experiment is conducted at the level of the individual, while boards are groups. Second, board members have much more experience and many more constraints they must attend to, while we use a relatively uninformed sample and

intentionally stripped away many of the complications of the board member selection process for simplicity.

In spite of these limitations, these studies collectively provide empirical evidence that group composition and group diversity decisions can be affected by threshold effects at the descriptive social norm. In our following studies, we provide additional experimental evidence directly testing our theoretical model to examine the influences of descriptive social norms, scrutiny, and visibility on group diversity decisions.

STUDY 2

In Study 2, we sought to test our theoretical model more directly by manipulating–rather than measuring–the descriptive social norms and scrutiny associated with the inclusion of females in a group. In addition, we sought to explore these phenomena in a new setting to establish their generalizability to groups besides corporate boards.

Study 2A: Group Diversity, Social Norms and Scrutiny

In Study 2A, we tested whether manipulating descriptive social norms and scrutiny affects decisions about whether to add a female candidate to a majority-male group with a sample of participants with work experience. Specifically, we investigated whether, as predicted in Hypothesis 2, individuals strive to meet descriptive social norms for diversity when under threat of possible scrutiny but not in cases where scrutiny is absent.

Method

Participants. Five hundred and fifty-six Master of Business Administration (MBA) students completed this study. This represented the entire incoming class at a U.S. business school. 57% of the participants were male, 25% had previous managerial experience before starting their MBA, and participants' average age was 27.7 years. Sample size was determined a priori, data analysis was conducted only once all data were collected, we do not exclude any data, and we report all measures and manipulations.

Procedures. Participants were asked to imagine their company had given them the task of assembling a seven-person panel for submission to an industry conference. They were told six of the seven panelists had already been determined, and they were responsible for selecting the final panelist. All participants saw an image of two women and four men representing the six predetermined panelists. Participants were randomly assigned to one of four experimental conditions (*surpassed social norm* or *unmet social norm* x *unscrutinized* or *scrutinized*) described below.

Participants saw images of five seven-person panels representing other panel submissions to the conference. Participants randomly assigned to the *surpassed social norm* condition saw that four of these other panel submissions had one woman each while one panel submission had no women on it (i.e. the average number of women on other panels was 0.8); participants randomly assigned to the *unmet social norm* condition saw that four of these panel submissions had three women each while one panel submission had two women on it (i.e. the average number of women on other panels was 2.8). Therefore, in the *surpassed social norm* condition, the participant's current panel (which

included two women) already exceeded the descriptive social norm for gender diversity (0.8 women); in the *unmet social norm* condition, the participant's current panel (which included two women) was below the descriptive social norm for gender diversity (2.8 women).

Participants were told panels were generally accepted based on speaker quality and years of industry experience of the panelists. Participants randomly assigned to the *unscrutinized* condition were told the review process was "blind": the names and photos of the panelists would not be submitted for evaluation (i.e. it would be impossible for the panels to be scrutinized with regards to gender composition). Participants randomly assigned to the *scrutinized* condition saw no such statement. Past research suggests that impression management concerns often arise when people simply know they are being evaluated (Leary & Kowalski, 1990), suggesting that when the evaluation process is not blind, scrutiny can be expected to affect decisions.[11]

Participants were then shown two potential candidates–Candidate A and Candidate B–for the final panelist. One image depicted a female candidate who had 10 years of industry experience and a speaker rating of 4.6; the other image depicted a male candidate who had 12 years of industry experience and a speaker rating of 4.8. Which

---

[11] In a separate pilot study, we asked participants to rate how much they agreed or disagreed on a 7-point scale with the statements, "My decision is under scrutiny with regards to the gender diversity of the panel" and "The reviewer will pay attention to the gender diversity of the panel when deciding which panels to accept". Participants in the *scrutinized* condition reported significantly higher scrutiny on gender diversity than participants in the *unscrutinized* condition ($M_{scrutinized} = 3.67$, $SD_{scrutinized} = 1.84$; $M_{unscrutinized} = 2.66$, $SD_{unscrutinized} = 1.92$, $t(150) = 3.34$, $p = 0.001$).

candidate was presented first as Candidate A (versus second as Candidate B) was randomized. Participants then rated their preference for the two candidates on a scale from 1 to 7 where 1 was labeled as "Strongly prefer Candidate A" and 7 was labeled as "Strongly prefer Candidate B." Study materials and a correlation matrix of all variables collected in this study are available in the *Online Supplement*.

Results

Consistent with Hypothesis 1b and as illustrated in Figure 5A, participants in the *scrutinized* condition had a significantly stronger preference for the female candidate in the *unmet social norm* condition than in the *surpassed social norm* condition ($t(277) = 2.24$; $p = 0.026$). In other words, participants whose diversity decisions could be scrutinized found it much more desirable to add a female candidate to a group when the group had not yet met the social norm for gender diversity compared to when the group had surpassed the social norm. However, consistent with Hypothesis 2, there were no differences in the preferences expressed for the female candidate between the *unmet social norm* and the *surpassed social norm* conditions when diversity decisions were not under scrutiny ($t(275) = 0.216$; $p = 0.829$).

Next, we checked if there was a significant interaction between surpassed social norms and the presence of scrutiny. We estimated a linear regression to predict the preference for the female candidate with indicators for our *scrutinized* condition, our *unmet social norm* condition, and the interaction between these two conditions (see Table 7, Model 1). The interaction term was positive but did not reach standard levels of

statistical significance ($p = 0.140$).[12] To strengthen our statistical power to detect an interaction, we conducted a follow-up study with incentivized decisions (note that we could not increase the sample size in this study because it already included every incoming MBA student at our selected university, and we were not able to incentivize the decisions of MBA students).

Study 2B: Replicating Study 2A With Incentives

In Study 2B, we sought to replicate our results from Study 2A but with real monetary stakes that would increase our statistical power to detect an interaction between the presence of scrutiny and a surpassed social norm for diversity. Again, we experimentally manipulated scrutiny and descriptive social norms to test for a causal relationship between these variables and the demographic characteristics of a newly-selected group member.

Method

Participants. Two hundred U.S. participants (51.5% male) were recruited through Amazon's Mechanical Turk and paid $0.15 to participate in a short online research study. Sample size was determined a priori, data analysis was conducted only once all data were collected, we do not exclude any data, and we report all measures and manipulations.

Procedures. As in Study 2A, participants were asked to imagine their company had given them the task of assembling a seven-person panel for submission to an industry

---

[12] We found a significant main effect of gender such that women had significantly higher preferences for the female candidate compared to men ($p = 0.022$), but there was no significant interaction.

conference, that six of the seven panelists had already been determined (two women and four men), and that they were responsible for selecting the final panelist. Again, they were randomly assigned to one of four experimental conditions.

In this study, we simplified the way the descriptive social norm was manipulated. Participants randomly assigned to the *surpassed social norm* condition were told competitive intelligence suggested the other panel submissions would have 1.25 women on average; participants randomly assigned to the *unmet social norm* condition were told the other panel submissions would have 2.75 women on average.

Participants were then told a reviewer would evaluate all panel submissions and choose to "accept" 75% of them. If their panel submission was accepted, participants would receive a bonus payment. All participants were initially allocated a $0.25 bonus, but participants had to "pay" to select the last panelist, and this cost was deducted from their promised bonus. Participants randomly assigned to the *unscrutinized* condition were told the review process was "blind": the names and photos of the panelists would not be submitted for evaluation (i.e. it would be impossible for the panels to be scrutinized with regards to gender composition). Participants randomly assigned to the *scrutinized* condition saw no such statement.

Participants were then offered the choice among three candidates for their final panelist. One image depicted a female candidate who had 10 years of industry experience, a speaker rating of 4.6, and cost $0.15 to select. The other images depicted male candidates who had similar qualifications (11 or 12 years of industry experience;

speaker ratings of 4.5 or 4.8) and cost $0.10 and $0.11 to select. Our outcome of interest was what fraction of participants in each condition selected the female candidate. We made the female candidate slightly more expensive to reflect research suggesting that women are more expensive to recruit and/or hire in contexts where diversity is lacking (e.g., on corporate boards and other contexts where less than 50% of the workforce is female, see Leslie, Manchester, & Dahm, 2016). Finally, participants reported their gender identity and whether they had ever attended or organized a conference in the past 10 years. Study materials and a correlation matrix of all variables collected in this study are available in the *Online Supplement*.

Results

Consistent with Hypothesis 1b, participants in the *scrutinized* condition were significantly more likely to select the female candidate in the *unmet social norm* condition than in the *surpassed social norm* condition ($z = 2.941$; $p = 0.0033$; see Figure 5B). Consistent with Hypothesis 2, there were no such differences in the likelihood of selecting the female candidate in the *unmet social norm* and the *surpassed social norm* conditions when there was no scrutiny ($z = 0.242$; $p = 0.808$).

To test for an interaction between the presence of scrutiny and unmet social norms, we estimated an ordinary least squares regression with robust standard errors to predict the choice of a female candidate with indicators for our *scrutinized* condition, our *unmet social norm* condition, and the interaction between these two conditions (see Table 7, Model 2). We found that the interaction term was positive and statistically significant

($\beta = 0.270$; $p = 0.028$). This further supports Hypothesis 2 that predicts when shaping the composition of groups, decision makers will only conform to the social norm for diversity when they are under scrutiny.

Discussion

Studies 2A and 2B directly manipulate scrutiny and descriptive social norms to provide direct tests of Hypotheses 1b and 2 and show that decision makers responsible for shaping group composition strive to increase group diversity when the group in question has not yet met the social norm for diversity on a scrutinized dimension (gender in the case of these studies). However, motivation to further increase diversity is reduced once the social norm has been met, and social norms do not exert this influence when scrutiny is not present.

STUDY 3: THE MODERATING EFFECT OF VISIBILITY

In Study 3, we manipulated descriptive social norms and a group's visibility to investigate whether the influence of descriptive social norms on decisions about group diversity is moderated by a group's visibility, and we also extend our study of group diversity to explore a social category besides gender.

Method

Participants. Six hundred and three U.S. participants (52.9% male; 80.4% Caucasian) were recruited through Amazon's Mechanical Turk to participate in a short online research study in exchange for $0.30. Sample size was determined a priori, data

analysis was conducted only once all data were collected, we do not exclude any data, and we report all measures and manipulations.

Procedures. Participants were told to imagine they were the manager of a team of five people and were hiring a sixth team member. All participants saw an image of one black man and four white men representing their current team. They were also told their HR department cared about the racial diversity of teams and the HR department could review team compositions and choose to punish teams deemed to have inadequate racial diversity, creating scrutiny on the dimension of racial diversity in all conditions. Participants were then randomly assigned to one of four experimental conditions. Participants randomly assigned to the *surpassed social norm* condition were told that other teams of their size included an average of 0.25 black people. Participants randomly assigned to the *unmet social norm* condition were told that other teams of their size included an average of 1.75 black people.

To manipulate visibility, participants were either randomly assigned to learn either: (1) their team was "not very important" in the company so there was a low probability that the HR department would review the composition of their team (the *low visibility* condition); or (2) their team was "very important" in the company so there was a high probability that the HR department would review the composition of their team (the *high visibility* condition).[13]

---

[13] In a separate pilot study, we asked participants to rate how much they agreed or disagreed on a 7-point scale with the statements, "My team is visible in the company"

Participants were then offered the choice of two candidates for their new team member. One image depicted a black male candidate who would come with a bonus of $0.03 to participants if they chose him; the other image depicted a white male candidate who would come with a bonus of $0.10 to participants if they chose him. We incentivized participants to choose the white man in order to overcome social desirability concerns and place some cost on increasing diversity. Participants were told they would keep the bonus associated with the candidate they chose unless the HR department reviewed their team and chose to penalize their team for a lack of racial diversity.

Finally, participants reported their racial and gender identities. Study materials and a correlation matrix of all variables collected in this study are available in the *Online Supplement*.

Results and Discussion

Consistent with Hypothesis 1b and all previous studies, participants were significantly more likely to select the black candidate in the *unmet social norm* condition than in the *surpassed social norm* condition ($z = 4.279$; $p < 0.0001$; see Figure 6). In other words, decision makers added the black candidate to their group at a lower rate once their group had surpassed the descriptive social norm for racial diversity. In addition, there was a significant main effect of visibility, such that participants were

---

and "My team receives a lot of attention in the company". Participants in the *high visibility* condition reported significantly higher scores on these items than participants in the *low visibility* condition ($M_{high\_visibility} = 6.39$, $SD_{high\_visibility} = 0.97$; $M_{low\_visibility} = 2.25$, $SD_{low\_visibility} = 1.49$; $t(147) = 20.04$; $p < 0.0001$).

significantly more likely to select the black candidate when their team was highly visible than when it was not ($z = 9.247$; $p < 0.0001$).

To test Hypothesis 3 that visibility moderates the effect of descriptive social norms, we tested for an interaction between visibility and social norms. To do this, we estimated an ordinary least squares regression with robust standard errors to predict the choice of the black candidate with indicators for our *high visibility* condition, our *unmet social norm* condition, and the interaction between these two conditions (see Table 8). Consistent with Hypothesis 3, we found the interaction term between visibility and norms was positive and statistically significant ($\beta = 0.151$; $p = 0.043$).

Overall, Study 3 conceptually replicates our previous studies, extends our findings to underrepresented groups besides women, and shows the moderating effect of visibility on decisions about group diversity.

GENERAL DISCUSSION

Across four experiments and one field study, we offer convergent evidence that those who shape the diversity of groups attend to and seek to conform to the descriptive social norms for diversity set by peer groups when under scrutiny. In Study 1, we showed that U.S. corporate boards are disproportionately likely to include exactly two women (the descriptive social norm), and they appear to lose motivation to add additional women once they have matched the descriptive social norm by including two female directors. We also found that these effects are more pronounced among more visible companies, consistent with our theory that these effects are driven in part by scrutiny and impression

management motives. In addition, we did not find any clustering when we analyzed data on the race/ethnicity of board members in our field data, consistent with our theory that scrutiny is required to produce clustering at the descriptive social norm.[14] In Studies 2 and 3, we directly manipulated descriptive social norms, scrutiny, and visibility to show that each of these influences group diversity decisions as our theory predicts in groups besides corporate boards and when we examine social categories besides gender.

Theoretical and Practical Implications

Our theory and findings help us understand how decision makers with the power to shape group composition respond to the threat of negative scrutiny surrounding diversity. Individuals responsible for group compositions look to descriptive social norms, matching the levels of diversity found in peer groups at an unusually high rate. This behavior leads to homogeneous levels of diversity across groups, providing another contributing explanation for the persistent underrepresentation of women and racial minorities in many organizational contexts. Our work also helps provide a fuller understanding of diversity-related hiring decisions, suggesting when women and racial minorities will be particularly attractive candidates for inclusion in groups and when groups can be expected to lessen their efforts to increase diversity.

Our findings suggest new avenues for policy makers seeking to increase diversity. Rather than simply targeting bias and stereotyping among those making hiring decisions

---

[14] As discussed in Study 1A, an analysis of media attention to board diversity showed that 97% of such articles discuss gender diversity, while just 18% even mention racial or ethnic diversity.

(e.g., through diversity training) or seeking to shape underrepresented candidates' preferences and skill sets (e.g., by training women to negotiate), more interventions may be needed to change the perceived norms around diversity. Groups appear to cluster at the descriptive social norm for diversity because it is an adaptive impression management strategy: by clustering at the social norm, they can escape negative scrutiny regarding their diversity levels. But the fact that groups can escape negative scrutiny once they reach the descriptive social norm for diversity implies that those scrutinizing these groups (e.g., shareholders, the media, etc.) may be too easily satisfied. Shifting the standards of those who scrutinize diversity as well as those of the decision makers capable of shaping group diversity from focusing on descriptive social norms in peer groups to instead achieving more ambitious norms (e.g., matching the levels of diversity in the general population) may be a promising new avenue for increasing the diversity of highly visible, scrutinized groups. If powerful institutions or individuals endorse new norms regarding gender and racial representation, perhaps this could lead to changes in the norms that influence group composition decisions (Paluck & Shepherd, 2012). For instance, decisions by the Supreme Court have been shown to change attitudes and perceptions of norms in the realm of gay rights (Tankard & Paluck, 2017).

Our work also points to scrutiny as a lever for change. Scrutiny can come from a variety of sources, but some sources may be more influential than others (Oliver, 1991). Applying greater scrutiny to group diversity should lead groups to increase their diversity. One extreme form of scrutiny when it comes to diversity is to enforce legal

penalties on public companies for a failure to diversify. However, even when policy makers have established laws mandating minimum levels of gender diversity on the corporate boards of public companies, some companies have elected to become private rather than comply with the laws (Miller, 2014). Forced compliance therefore comes with the risk of creating at least some reactance (Dobbin, Schrage, & Kalev, 2015). An alternative to mandated diversity may be to shower positive attention on groups that reach high levels of diversity. Treating diversity as an ideal may help reshape perceptions of the relevant norm, leading injunctive norms (or norms about ideals) to overshadow descriptive social norms.

Limitations and Future Research

One paradox suggested by our theorizing and empirics surrounds changing descriptive norms: U.S. corporate boards shifted from clustering at one woman to clustering at two women (albeit slowly) over the last twenty years in spite of the fact that our theorizing about diversity thresholds would predict a stagnation of board diversity at the one-woman threshold. A noteworthy fact, however, is that this shift in clustering followed the passage of Norway's "Women on Boards" act in 2003. This legislation required public and state-owned companies in Norway to include at least 40% women and may have made the topic of gender diversity on corporate boards in the U.S. more salient at that time, providing increased scrutiny of boards with few women and making the need for gender diversity more salient, driving the shift to twokenism from tokenism.

Future research exploring how descriptive social norms can be shifted in the context of diversity would be extremely valuable.

Another puzzling question raised by our findings is whether more diverse groups may actually discriminate more than less diverse groups. We cannot evaluate whether any specific group or organization is actively "managing" diversity for impression management reasons. However, overall, we do see a pattern suggesting this is the case and suggesting that–contrary to the expectation that more diverse groups will attract more women and racial minority candidates (Avery, 2003; Avery & McKay, 2006)–such groups are less likely to select women and racial minorities than others after reaching the descriptive social norm for diversity. It would be valuable for future research to examine when and how social norms around diversity can hurt rather than help women and minorities.

Although our field and experimental studies provide convergent evidence in support of our theory and hypotheses, in our experiments, we only examine the judgments and decisions of individuals, while group member selection processes are varied and complex and often involve many decision makers. Extensive past research has shown that studies of individual decisions and insights about individual psychology can further our understanding of group and organizational outcomes (Greve, 2008; Highhouse, Brooks, & Gregarus, 2009; Simon & Houghton, 2003; Staw, 1991). However, there are unquestionably limitations in our approach.

We only test our theorizing in a single field setting (albeit in an economically and organizationally important one). Future research examining how these phenomena play out in other important organizational contexts would undoubtedly be useful. Our experiments may also be susceptible to demand effects, which could limit their external validity. In addition, in our field setting and in our experiments, the groups we examine are relatively small in size (i.e. less than 20 members). Additional research in exploring how group size moderates the effects of descriptive social norms and scrutiny could be informative. For example, in larger groups, the behavior of peer groups could feel less relevant as the size of the group might create a greater sense of its uniqueness, reducing pressure to conform to descriptive social norms. Alternatively, larger groups may feel more scrutinized because of their size, leading them to react more dramatically to descriptive social norms.

Finally, more research into the psychological mechanisms that lead descriptive social norms and scrutiny to produce the group diversity threshold effects we document could be illuminating. Past research suggests that norms may be particularly relevant in the context of group diversity decisions because of ambiguity about how much diversity is enough and the fear of being singled out from peers (Ahmadjian & Robinson, 2001; Festinger, 1954; Sherif, 1936; Zavyalova et al., 2012). Future research isolating the specific mechanisms through which descriptive social norms exert their influence would be valuable and could help identify potent interventions for changing salient norms.

Future research testing new interventions to reduce the reliance on descriptive social norms and make other norms more salient would also be extremely valuable.

Conclusion

Our work highlights the important roles that descriptive social norms, goal setting, scrutiny, and visibility play in shaping decisions about group diversity while answering questions about how individuals assess whether a group is diverse and how groups respond to scrutiny around their diversity levels. We find empirical evidence that descriptive social norms and threshold effects lead to an overabundance of groups with exactly the same level of diversity in an important organizational context, providing evidence of a previously unexplored phenomenon that may contribute to the underrepresentation of women and minorities in many organizational groups. By shedding light on novel factors that influence group diversity decisions, we illuminate potential new avenues for increasing the diversity of groups.

REFERENCES

Adams, R. B., & Ferreira, D. (2009). Women in the boardroom and their impact on governance and performance. *Journal of Financial Economics*, *94*(2), 291–309.

Aguilera, R. V., Rupp, D. E., Williams, C. A., & Ganapathi, J. (2007). Putting the S back in corporate social responsibility: A multilevel theory of social change in organizations. *Academy of Management Review*, *32*(3), 836–863.

Ahmadjian, C. L., & Robinson, P. (2001). Safety in numbers: Downsizing and the deinstitutionalization of permanent employment in Japan. *Administrative Science Quarterly*, *46*(4), 622–654.

Allen, E. J., Dechow, P. M., Pope, D. G., & Wu, G. (2016). Reference-dependent preferences: Evidence from marathon runners. *Management Science*.

Angrist, J. D., & Pischke, J.-S. (2008). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Avery, D. R. (2003). Reactions to diversity in recruitment advertising--are differences black and white? *Journal of Applied Psychology*, *88*(4), 672.

Avery, D. R., & McKay, P. F. (2006). Target practice: An organizational impression management approach to attracting minority and female job applicants. *Personnel Psychology*, *59*(1), 157–187.

Bainbridge, S. M., & Henderson, M. T. (2014). Boards-R-Us: reconceptualizing corporate boards. *Stan. L. Rev.*, *66*, 1051.

Bell, J. M., & Hartmann, D. (2007). Diversity in everyday discourse: The cultural ambiguities and consequences of "happy talk." *American Sociological Review*, *72*(6), 895–914.

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, *34*(6), 1080–1109.

Brammer, S., & Millington, A. (2006). Firm size, organizational visibility and corporate philanthropy: An empirical analysis. *Business Ethics: A European Review*, *15*(1), 6–18.

Brands, R. A., & Fernandez-Mateo, I. (2016). Leaning Out: How Negative Recruitment Experiences Shape Women's Decisions to Compete for Executive Roles. *Administrative Science Quarterly*, 0001839216682728.

Buckley, C. (2016, January 15). Another Oscar Year, Another All-White Ballot. *The New York Times*. Retrieved from https://www.nytimes.com/2016/01/16/movies/oscar-ballot-is-all-white-for-another-year.html

Campbell, J. L. (2007). Why would corporations behave in socially responsible ways? An institutional theory of corporate social responsibility. *Academy of Management Review*, *32*(3), 946–967.

Castilla, E. J., & Benard, S. (2010). The paradox of meritocracy in organizations. *Administrative Science Quarterly*, *55*(4), 543–676.

Chiu, S.-C., & Sharfman, M. (2011). Legitimacy, visibility, and the antecedents of corporate social performance: An investigation of the instrumental perspective. *Journal of Management*, *37*(6), 1558–1585.

Cialdini, R. B. (2003). Crafting normative messages to protect the environment. *Current Directions in Psychological Science*, *12*(4), 105–109.

Cialdini, R. B. (2007). Descriptive social norms as underappreciated sources of social control. *Psychometrika*, *72*(2), 263.

Cialdini, R. B., Kallgren, C. A., & Reno, R. R. (1991). A focus theory of normative conduct: A theoretical refinement and reevaluation of the role of norms in human behavior. *Advances in Experimental Social Psychology*, *24*, 201–234.

Cialdini, R. B., Reno, R. R., & Kallgren, C. A. (1990). A focus theory of normative conduct: recycling the concept of norms to reduce littering in public places. *Journal of Personality and Social Psychology*, *58*(6), 1015.

Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity and compliance. Retrieved from http://psycnet.apa.org/psycinfo/1998-07091-021

Coffman, L. C., Featherstone, C. R., & Kessler, J. B. (2014). Can Social Information Affect What Job You Choose and Keep? *American Economic Journal: Applied Economics*.

Desai, V. M. (2011). Mass media and massive failures: Determining organizational efforts to defend field legitimacy following crises. *Academy of Management Journal*, *54*(2), 263–278.

Dezső, C. L., Ross, D. G., & Uribe, J. (2016). Is there an implicit quota on women in top management? A large-sample statistical analysis. *Strategic Management Journal*, *37*(1), 98–115.

DiMaggio, P., & Powell, W. W. (1983). The iron cage revisited: Collective rationality and institutional isomorphism in organizational fields. *American Sociological Review*, *48*(2), 147–160.

Dobbin, F., Schrage, D., & Kalev, A. (2015). Rage against the iron cage: The varied effects of bureaucratic personnel reforms on diversity. *American Sociological Review*, *80*(5), 1014–1044.

Elsbach, K. D., & Sutton, R. I. (1992). Acquiring organizational legitimacy through illegitimate actions: A marriage of institutional and impression management theories. *Academy of Management Journal*, *35*(4), 699–738.

Elsbach, K. D., Sutton, R. I., & Principe, K. E. (1998). Averting expected challenges through anticipatory impression management: A study of hospital billing. *Organization Science*, *9*(1), 68–86.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, *7*(2), 117–140.

Fombrun, C. (1996). *Reputation*. Wiley Online Library.

Fombrun, C., & Shanley, M. (1990). What's in a name? Reputation building and corporate strategy. *Academy of Management Journal*, *33*(2), 233–258.

Forbes, D. P., & Milliken, F. J. (1999). Cognition and corporate governance: Understanding boards of directors as strategic decision-making groups. *Academy of Management Review*, *24*(3), 489–505.

Gardberg, N. A., & Fombrun, C. J. (2006). Corporate citizenship: Creating intangible assets across institutional environments. *Academy of Management Review*, *31*(2), 329–346.

Gino, F., & Pierce, L. (2010). Robin Hood under the hood: Wealth-based discrimination in illicit customer help. *Organization Science*, *21*(6), 1176–1194.

Goldstein, N. J., Cialdini, R. B., & Griskevicius, V. (2008). A room with a viewpoint: Using social norms to motivate environmental conservation in hotels. *Journal of Consumer Research*, *35*(3), 472–482.

Greve, H. R. (2008). A behavioral theory of firm growth: Sequential attention to size and performance goals. *Academy of Management Journal*, *51*(3), 476–494.

Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology*, *38*(1), 79–109.

Highhouse, S., Brooks, M. E., & Gregarus, G. (2009). An organizational impression management perspective on the formation of corporate reputations. *Journal of Management*, *35*(6), 1481–1493.

Kallgren, C. A., Reno, R. R., & Cialdini, R. B. (2000). A focus theory of normative conduct: When norms do and do not affect behavior. *Personality and Social Psychology Bulletin*, *26*(8), 1002–1012.

Kanter, R. M. (1977). Some effects of proportions on group life: Skewed sex ratios and responses to token women. *American Journal of Sociology*, 965–990.

King, B. G. (2008). A political mediation model of corporate response to social movement activism. *Administrative Science Quarterly*, *53*(3), 395–421.

Konrad, A. M., Kramer, V., & Erkut, S. (2008). Critical Mass: The Impact of Three or More Women on Corporate Boards. *Organizational Dynamics*, *37*(2), 145–164.

Leary, M. R., & Kowalski, R. M. (1990). Impression management: A literature review and two-component model. *Psychological Bulletin*, *107*(1), 34.

Lee, J. C. (2017, January 13). Trump's Cabinet So Far Is More White and Male Than Any First Cabinet Since Reagan's. *The New York Times*. Retrieved from https://www.nytimes.com/interactive/2017/01/13/us/politics/trump-cabinet-women-minorities.html

Leslie, L. M., Manchester, C., & Dahm, P. (2017). Why and When Does the Gender Gap Reverse? Diversity Goals and the Pay Premium for High Potential Women. *Academy of Management Journal*, *60*(2), 402–432.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705.

Matsa, D. A., & Miller, A. R. (2013). A female style in corporate leadership? Evidence from quotas. *American Economic Journal: Applied Economics*, *5*(3), 136–169.

McDonnell, M. H., & King, B. (2013). Keeping up appearances reputational threat and impression management after social movement boycotts. *Administrative Science Quarterly*, *58*(3), 387–419.

Merchant, N. (2013, October 7). Viewpoint: Twitter's All-Male Board Spells Failure. *Time*. Retrieved from http://ideas.time.com/2013/10/07/viewpoint-twitters-all-male-board-spells-failure/

Miller, C. C. (2013, October 4). Curtain is rising on a tech premiere with (as usual) a mostly male cast. *The New York Times*. Retrieved from http://www.nytimes.com/2013/10/05/technology/as-tech-start-ups-surge-ahead-women-seem-to-be-left-behind.html

Miller, C. C. (2014, June 19). Women on the Board: Quotas Have Limited Success. *The New York Times*. Retrieved from https://www.nytimes.com/2014/06/20/upshot/women-on-the-board-quotas-have-limited-success.html

Moskowitz, T., & Wertheim, L. J. (2011). *Scorecasting: The hidden influences behind how sports are played and games are won*. Crown Archetype. Retrieved from https://books.google.com/books?hl=en&lr=&id=0lSi_eCQvNwC&oi=fnd&pg=PA1&dq=moskowitz+wertheim+scorecasting&ots=Egl8R3BtiF&sig=E3TaC8aweoucW_5RyN6As6R5fCI

Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin*, *34*(7), 913–923.

Oliver, C. (1991). Strategic responses to institutional processes. *Academy of Management Review*, *16*(1), 145–179.

Paluck, E. L., & Shepherd, H. (2012). The salience of social referents: A field experiment on collective norms and harassment behavior in a school social network. *Journal of Personality and Social Psychology*, *103*(6), 899.

Pierce, L., Snow, D. C., & McAfee, A. (2015). Cleaning house: The impact of information technology monitoring on employee theft and productivity. *Management Science*, *61*(10), 2299–2319.

Pollock, T. G., & Rindova, V. P. (2003). Media legitimation effects in the market for initial public offerings. *Academy of Management Journal*, *46*(5), 631–642.

Pope, D., & Simonsohn, U. (2011). Round numbers as goals evidence from baseball, SAT takers, and the lab. *Psychological Science*, *22*(1), 71–79.

Prentice, D. A., & Miller, D. T. (1993). Pluralistic ignorance and alcohol use on campus: some consequences of misperceiving the social norm. *Journal of Personality and Social Psychology*, *64*(2), 243.

Rubinstein, R. Y., & Kroese, D. P. (2011). *Simulation and the Monte Carlo method* (Vol. 707). John Wiley & Sons. Retrieved from https://books.google.com/books?hl=en&lr=&id=yWcvT80gQK4C&oi=fnd&pg=

PT10&dq=rubinstein+kroese&ots=xosXA8fboK&sig=Yk6woEUxf0PHvWY5QDk7vauSVJ4

Ryan, P. (2016, February 2). #OscarsSoWhite controversy: What you need to know. *USA Today*. Retrieved from https://www.usatoday.com/story/life/movies/2016/02/02/oscars-academy-award-nominations-diversity/79645542/

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, *18*(5), 429–434.

Shaffer, L. S. (1983). Toward Pepitone's vision of a normative social psychology: What is a social norm? *Journal of Mind and Behavior*.

Shemla, M., Meyer, B., Greer, L., & Jehn, K. A. (2014). A review of perceived diversity in teams: Does how members perceive their team's composition affect team processes and outcomes? *Journal of Organizational Behavior*.

Sherif, M. (1936). The psychology of social norms.

Simon, M., & Houghton, S. M. (2003). The relationship between overconfidence and the introduction of risky products: Evidence from a field study. *Academy of Management Journal*, *46*(2), 139–149.

Smale, A., & Miller, C. C. (2015, March 6). Germany Sets Gender Quota in Boardrooms. *The New York Times*. Retrieved from

https://www.nytimes.com/2015/03/07/world/europe/german-law-requires-more-women-on-corporate-boards.html

S&P Dow Jones Indices. (2015). *S&P U.S. Indices Methodology, October 2015*. Retrieved from https://us.spindices.com/indices/equity/sp-500

Staats, B. R., Dai, H., Hofmann, D., & Milkman, K. L. (2016). Motivating process compliance through individual electronic monitoring: An empirical examination of hand hygiene in healthcare. *Management Science, 63*(5), 1563–1585.

Staw, B. M. (1991). Dressing up like an organization: When psychological theories can explain organizational action. *Journal of Management*, *17*(4), 805–819.

Sutton, R. I., & Galunic, D. C. (1996). Consequences of public scrutiny for leaders and their organizational image and its management. *Research in Organizational Behavior*, *18*, 201–250.

Swim, J. K., & Hyers, L. L. (1999). Excuse me—What did you just say?!: Women's public and private responses to sexist remarks. *Journal of Experimental Social Psychology*, *35*(1), 68–88.

Tankard, M. E., & Paluck, E. L. (2017). The effect of a Supreme Court decision regarding gay marriage on social norms and personal attitudes. *Psychological Science*, *28*(9), 1334–1344.

Torchia, M., Calabrò, A., & Huse, M. (2011). Women directors on corporate boards: From tokenism to critical mass. *Journal of Business Ethics*, *102*(2), 299–317.

Totenberg, N. (2016, September 23). Donald Trump Unveils New, More Diverse Supreme Court Short List. *NPR*. Retrieved from http://www.npr.org/2016/09/23/495216645/donald-trump-unveils-new-more-diverse-supreme-court-short-list

Unzueta, M. M., & Binning, K. R. (2010). Which racial groups are associated with diversity? *Cultural Diversity and Ethnic Minority Psychology*, *16*(3), 443–446.

Unzueta, M. M., & Binning, K. R. (2012). Diversity Is in the Eye of the Beholder How Concern for the In-Group Affects Perceptions of Racial Diversity. *Personality and Social Psychology Bulletin*, *38*(1), 26–38.

Unzueta, M. M., Knowles, E. D., & Ho, G. C. (2012). Diversity Is What You Want It to Be How Social-Dominance Motives Affect Construals of Diversity. *Psychological Science*, *23*(3), 303–309.

Zavyalova, A., Pfarrer, M. D., Reger, R. K., & Shapiro, D. L. (2012). Managing the message: The effects of firm actions and industry spillovers on media coverage following wrongdoing. *Academy of Management Journal*, *55*(5), 1079–1101.

FIGURES

Figure 1

Comparison of Actual Distribution of Women on (A) S&P 1500 Boards and (B) S&P 500 Boards with Simulated Expected Distribution of Women

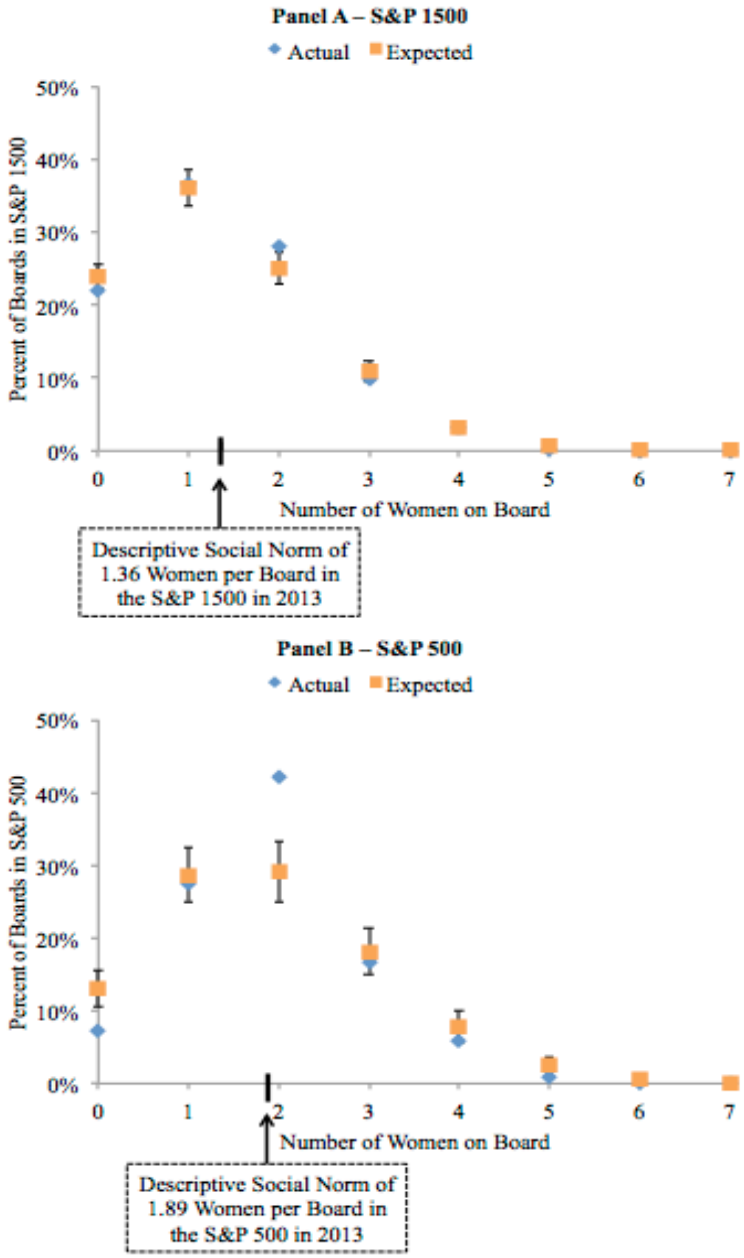
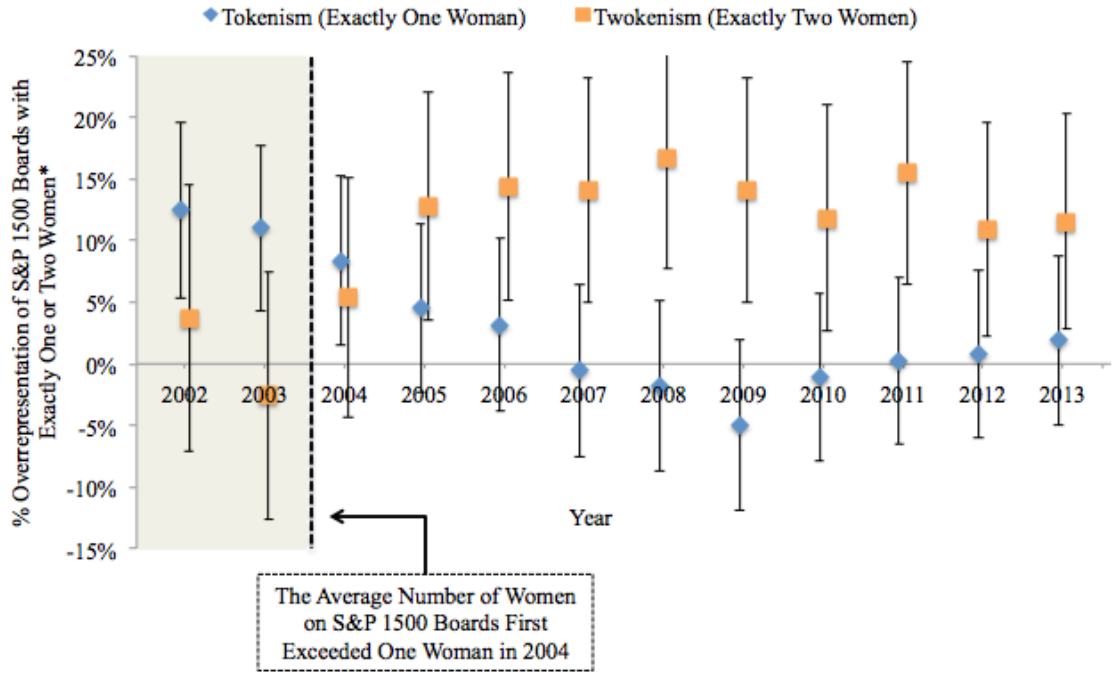


Figure 2

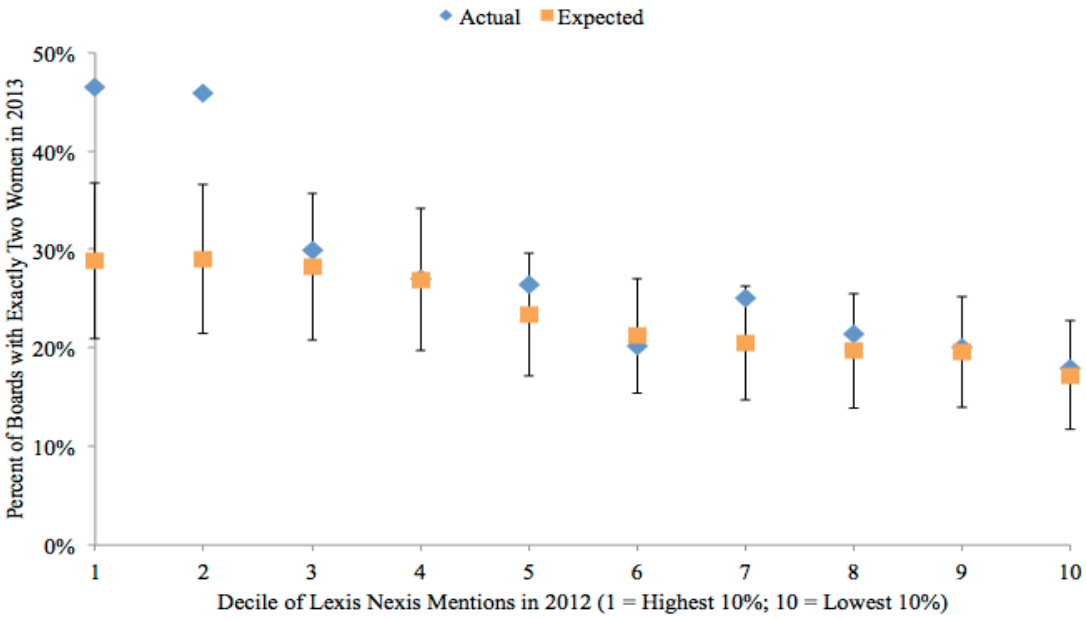
How Tokenism and Twokenism Shifted as Social Norms Changed from 2002 to 2013



* % Overrepresentation of S&P 1500 Boards with Exactly One Woman is calculated for each year using the following formula: [actual number of S&P 1500 boards with exactly one woman – expected number of S&P 1500 boards with exactly one woman] / [expected number of S&P 1500 boards with exactly one woman] where the expected number is calculated using our standard Monte Carlo simulation method. % Overrepresentation of S&P 1500 Boards with Exactly Two Women is calculated for each year using the following formula: [actual number of S&P 1500 boards with exactly two women – expected number of S&P 1500 boards with exactly two women] / [expected number of S&P 1500 boards with exactly two women] where the expected number is calculated using our standard Monte Carlo simulation method.

Figure 3

Firm Visibility Moderates the Extent of Twokenism



Note: This figure depicts the comparison between the actual percent of boards with exactly two women and the simulated expected percent of boards with exactly two women in 2013 as moderated by a firm's visibility. Firm visibility is operationalized using media mentions measured from Lexis Nexis, and firms are grouped into deciles based on the number of media mentions they received in 2012.

Figure 4

Participants in Study 1C Less Likely to Increase Gender Diversity of Boards Once Boards Include Two Women (and Thus Exceed the Social Norm)

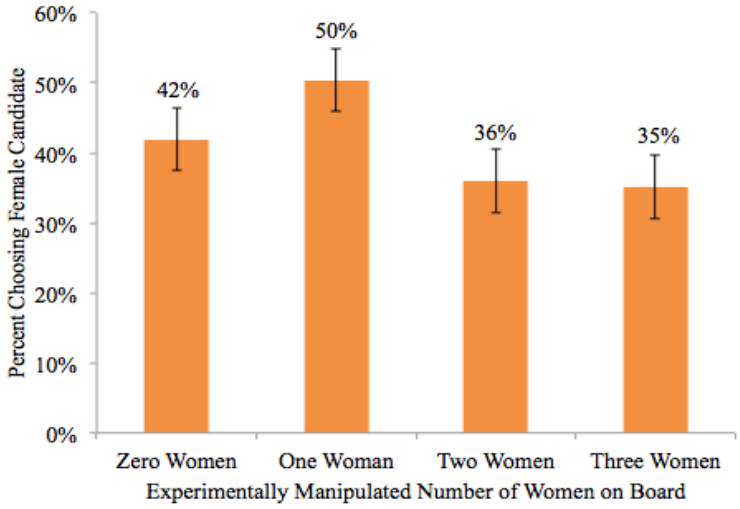


Figure 5A

Participants' Preferences for Women Are Influenced by Social Norms and Scrutiny in Study 2A

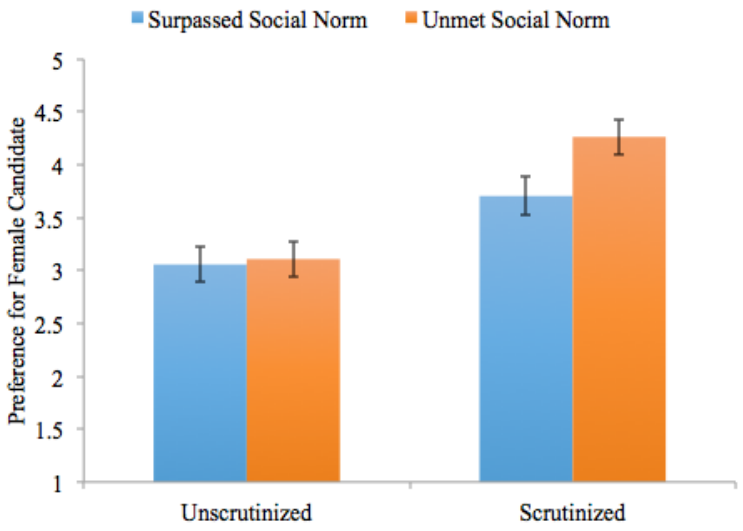


Figure 5B

Interaction Between Social Norms and Scrutiny in Study 2B

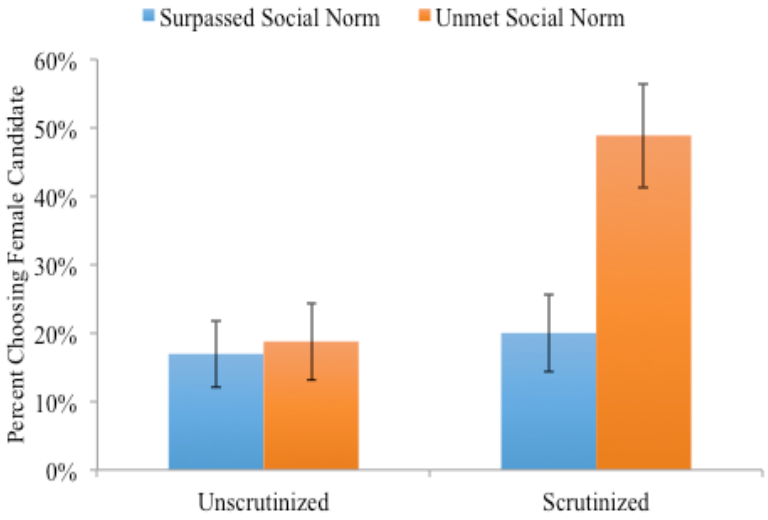
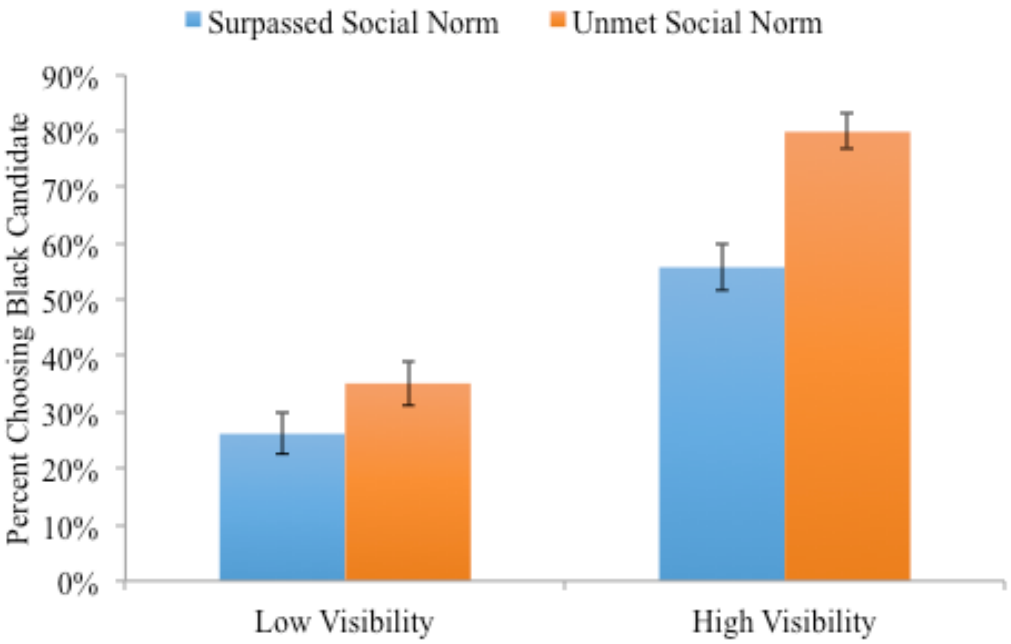


Figure 6

Interaction Between Social Norms and Visibility in Study 3



TABLES

Table 1

Summary Statistics Describing S&P 1500 Dataset

	Proportion of all Directors
Male	86%
Female	14%
Caucasian	91%
Asian	3.0%
Black	3.7%
Hispanic	1.7%
Other Ethnicity	0.81%
1 Board Seat	84%
2 Board Seats	13%
3 Board Seats	2.8%
4 Board Seats	0.37%
5 Board Seats	0.07%

Table 2

Correlation Matrix for S&P 1500 Board Data in 2013 (*N* = 1,441)

	1	2	3	4	5	6
1. Size of Board	1.00					
2. Number of Female Directors	0.51***	1.00				
3. Number of Racial Minority Directors	0.36***	0.30***	1.00			
4. Logarithm of Market Capitalization	0.44***	0.36***	0.31***	1.00		
5. Logarithm of Media Mentions	0.43***	0.38***	0.32***	0.59***	1.00	
6. Member of S&P 500	0.43***	0.36***	0.29***	0.71***	0.59***	1.00

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively

Table 3

Comparison of Actual and Expected Number of Female Directors Across S&P 1500 Boards of Different Sizes

Size of Board	*n*	Excess Percentage of Boards Observed with 0 Female Directors	Excess Percentage of Boards Observed with 1 Female Director	Excess Percentage of Boards Observed with 2 Female Directors	Excess Percentage of Boards Observed with 3 Female Directors
6 or fewer	124	2.74% (2.89%)	-5.79% (11.23%)	-16.96% (28.15%)	37.36% (109.43%)
7	199	-2.60% (3.34%)	4.38% (8.98%)	10.06% (16.79%)	-66.06% (54.18%)
8	241	-15.98%** (5.57%)	23.15%** (8.04%)	-8.42% (11.18%)	-24.81% (24.50%)
9	283	-26.32%** (8.05%)	14.16%* (7.22%)	20.60%* (9.09%)	-34.31%* (14.04%)
10	235	-38.00%*** (10.85%)	18.88%* (8.58%)	16.51% (10.14%)	-16.12% (13.57%)
11	198	-50.56%*** (14.71%)	6.67% (9.95%)	28.69%** (10.95%)	7.47% (13.42%)
12	100	-76.36%** (28.08%)	-22.97% (15.09%)	64.40%*** (15.38%)	-6.40% (17.34%)
13 or more	134	-58.37%* (24.01%)	-20.70% (13.63%)	49.34%*** (13.10%)	22.85% (15.55%)

Note: This table reports the difference between the actual percent of boards with a given number of female directors and the simulated expected percent of boards with that number of female directors conditional on the size of the board. Standard deviations are reported in parentheses.

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively

Table 4

Boards Less Likely to Add Additional Women Once They Include at Least Two Women

	Board Added Woman = 1 (ordinary least squares regression)			
Sample:	**Model 1 S&P 1500**	**Model 2 S&P 1500**	**Model 3 S&P 500**	**Model 4 S&P 500**
Number of Women on Board	-0.0033	-0.039***	-0.0056	-0.035*
	(0.0079)	(0.0090)	(0.012)	(0.015)
Indicator for Two or More Women on Board	-0.039*	-0.034*	-0.092***	-0.090***
	(0.016)	(0.016)	(0.023)	(0.024)
Controls Present	No	Yes	No	Yes
Observations	9,989	9,936	4,131	4,117
R^2	0.0032	0.030	0.017	0.045

Note. This table shows a series of ordinary least squares (OLS) regressions predicting whether boards add women conditional on the number of women already on the board and whether the board had met the descriptive social norm for gender diversity (i.e. already had at least two women) in the S&P 1500 (Models 1 and 2) and the S&P 500 (a subset of the S&P 1500; Models 3 and 4). Robust standard errors are in parentheses. When controls are present, regressions include fixed effects for board size, fixed effects for industry, fixed effects for stock market index, and a continuous control for market capitalization.

*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively

Table 5

More Visible Companies Show Larger Discontinuities at the Descriptive Social Norm

	Board Added Woman = 1 (ordinary least squares regression)	
	Model 1	**Model 2**
Number of Women on Board	-0.017*	-0.043***
	(0.0083)	(0.0090)
Indicator for Two or More Women on Board	-0.27	-0.023
	(0.017)	(0.017)
Centered Logarithm of Media Mentions	0.026***	0.017**
	(0.0042)	(0.0049)
Number of Women on Board x Centered Logarithm of Media Mentions	-0.0018	-0.00018
	(0.0049)	(0.0053)
Indicator for Two or More Women on Board x Centered Logarithm of Media Mentions	-0.021*	-0.021*
	(0.010)	(0.010)
Controls Present	No	Yes
Observations	9,781	9,743
R^2	0.012	0.033

Note. This table shows two ordinary least squares (OLS) regressions predicting whether boards add women conditional on the number of women already on the board and whether the board had met the descriptive social norm for gender diversity (i.e. already had at least two women), interacted with the centered logarithm of the number of media mentions a company receives. Robust standard errors are in parentheses. When controls are present, regressions include fixed effects for board size, fixed effects for industry, fixed effects for stock market index, and a continuous control for market capitalization.
*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively

Table 6

Regression Predicting the Selection of the Female Candidate to Serve on a Corporate Board in Study 1C

	B
Number of Women on Original Board	0.0401 (0.0448)
Original Board Has Two or More Women	-0.187† (0.100)
Observations	479
R^2	0.0134

Note. This table shows the results of an ordinary least squares regression (OLS) predicting whether participants added a woman to a board conditional on the number of women already on the board and whether the board had met the descriptive social norm for gender diversity (i.e. already included at least two women). Robust standard errors are in parentheses.
†, *, **, and *** denote significance at the 10%, 5%, 1%, and 0.1% levels, respectively

Table 7

Regression Predicting Preference for Female Candidates to Serve on Panels in Studies 2A and 2B

	Model 1, Study 2A	**Model 2, Study 2B**
DV:	***Rating of Female Candidate***	***Chose Female Candidate***
Scrutinized	0.647**	0.031
	(0.243)	(0.075)
Unmet Social Norm	0.0511	0.018
	(0.244)	(0.075)
Scrutinized x Unmet Social Norm	0.508	0.270*
	(0.344)	(0.122)
Observations	556	200
R^2	0.0563	0.0837

Note. These ordinary least squares (OLS) regressions present the preference for the female candidate to serve on a panel in Studies 2A and 2B. Scrutinized is an indicator for the Scrutinized condition. Unmet Social Norm is an indicator for the unmet social norm condition. Robust standard errors are in parentheses.
*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively

Table 8

Regression Predicting the Selection of a Black Candidate for a Team in Study 3

	B
High Visibility	0.296***
	(0.055)
Unmet Social Norm	0.089
	(0.053)
High Visibility x Unmet Social Norm	0.151*
	(0.043)
Observations	603
R^2	0.175

Note. This ordinary least squares (OLS) regression predicts whether participants chose the black candidate to serve on a team in Study 3. High Visibility is an indicator for the High Visibility condition. Unmet Social Norm is an indicator for the unmet social norm condition. Robust standard errors are in parentheses.
*, **, and *** denote significance at the 5%, 1%, and 0.1% levels, respectively

CHAPTER 2. THE ISOLATED CHOICE EFFECT AND ITS IMPLICATIONS FOR GENDER DIVERSITY IN ORGANIZATIONS

Edward H. Chang, Erika L. Kirgios, Aneesh Rai, and Katherine L. Milkman

Forthcoming in *Management Science*

ABSTRACT

We highlight a feature of personnel selection decisions that can influence the gender diversity of groups and teams. Specifically, we show that people are less likely to choose candidates whose gender would increase group diversity when making personnel selections in isolation (i.e., when they are responsible for selecting a single group member) than when making collections of choices (i.e., when they are responsible for selecting multiple group members). We call this the *isolated choice effect*. Across 6 preregistered experiments (n=3,509), we demonstrate that the isolated choice effect has important consequences for group diversity. When making sets of hiring and selection decisions (as opposed to making a single hire), people construct more gender-diverse groups. Mediation and moderation studies suggest that people do not attend as much to diversity when making isolated selection choices, which drives this effect.

Link to OSF with Online Supplement and open data:
https://osf.io/p2c8h/?view_only=96cb464b1b674750a3579a6228a89bdd

INTRODUCTION

Many organizations publicly espouse commitments to increase their diversity and inclusiveness. For example, the majority of Fortune 500 companies boast executives responsible for workplace diversity initiatives (Kwoh, 2012) and include diversity statements on their websites (Jones & Donnelly, 2017). In spite of this, many well-intentioned organizations remain remarkably homogeneous, which has prompted a large body of research exploring why homogeneity persists and tactics for increasing workplace diversity (Babcock, Recalde, Vesterlund, & Weingart, 2017; Bohnet, Van Geen, & Bazerman, 2015; Schroeder & Risen, 2016).

Notably, diversity objectives are typically set by organizational leaders who have a bird's-eye view of how hiring decisions shape the overall diversity of their organizations. However, the implementation of these objectives is often left to individual managers and teams who frequently make hiring decisions one at a time, making it challenging to take a global perspective on the impact each hire will have on organizational diversity. We propose that this common feature of hiring decisions–that they are often made in isolation–may have important and previously unappreciated implications.

In this paper, we show that people are less likely to choose candidates whose gender would increase group diversity when making personnel selections in isolation (i.e., when they are responsible for selecting only a single group member) than when making collections of choices (i.e., when they are responsible for selecting multiple

group members). We call this phenomenon the *isolated choice effect*, and it means groups constructed through an aggregation of isolated selection decisions are less diverse than groups whose members are selected in collections.

When people make hiring or selection decisions in isolation, we theorize that they attend less to how their selected candidate will affect the diversity of the group than when making collections of such decisions. Past research has shown that people are able to rapidly and accurately form impressions of the diversity of a group of people (Phillips, Slepian, & Hughes, 2018). The diversity of a collection of people selected together is therefore likely to be easy to assess and salient. However, because diversity is inherently a group-level property (Harrison & Klein, 2007) and any one individual cannot be "diverse," diversity is less tangible and therefore likely to be less salient when making selection decisions in isolation. Past research has shown that salient attributes are over-weighted when we make choices (Bordalo, Gennaioli, & Shleifer, 2012, 2013). We therefore hypothesize that the decreased salience of diversity when selection decisions are made in isolation (instead of in collections) will produce the isolated choice effect.

Past research on choice bracketing in consumer settings provides suggestive evidence that making isolated selection decisions may lessen the gender diversity of constructed groups (Read, Loewenstein, & Rabin, 1999). Specifically, people have been shown to select less variety in products when making consumption decisions one at a time (e.g., on a series of separate occasions) rather than simultaneously (Read, Antonides, Van den Ouden, & Trienekens, 2001; Read & Loewenstein, 1995; Simonson, 1990;

Simonson & Winer, 1992). In one canonical study, Simonson (1990) offered students snack choices at three meetings spread across three weeks. Students who were randomized to choose one snack each week were significantly more likely to choose the same snack each time (and thus a less diverse set of snacks) than students randomized to simultaneously choose snacks for all three meetings at the first weekly gathering.[15]

Notably, there are important differences between consumer choice and hiring decisions that make it unclear whether we should observe the same patterns in both settings. First, past research on choice bracketing in the consumer space typically focuses on consumers' preferences for "variety," wherein consumers choose different items (rather than the exact same item) in consumption bundles. For example, choosing a Snickers and a Twix would be considered variety-seeking, even though both are chocolate candy bars, as this represents more variety than choosing two Snickers bars (or two Twix bars). Here, the exact same consumer product can be consumed more than once. For hiring or selection decisions, because the same person cannot be hired repeatedly, all sets of decisions tautologically have identical variety since each hiring decision requires selecting a different person. However, some sets of decisions differ in their demographic "diversity," which is the focus of this paper.

---

[15] More broadly, the isolated choice effect is related to extensive prior work on the behavioral consequences of narrow versus broad decision frames in a wide range of settings ranging from labor supply decisions to budgeting (Camerer, Babcock, Loewenstein, & Thaler, 1997; Thaler, 1999).

A second key distinction between past research on choice bracketing in consumer contexts and our work on personnel selection pertains to what drives the underlying decisions. Many of the mechanisms shown to operate in the realm of consumer choice cannot apply in the realm of personnel selection. For instance, past research has posited that one reason choosing products all at once rather than over time leads to product diversification is because people have uncertain forecasts of their future preferences (Read & Loewenstein, 1995; Simonson, 1990). Further, because some choices made in sets are made for future consumption periods (while isolated choices are made at the moment of consumption), people may overestimate how much their tastes will change or how satiated they will be by repeatedly consuming the same product, or they may choose variety to reduce the risk of consuming something undesirable repeatedly. In personnel selection decisions, however, these factors are unlikely to play a role because the same person cannot be hired for multiple positions, and people are not "consumed" like products. While expecting to get bored with Twix bars quickly or worrying that you won't like Twix bars as much as anticipated may prompt the inclusion of a Snickers bars in your consumption bundle, a hiring manager is unlikely to think about choosing male and female job candidates in this way. Every job candidate is unique (i.e., not all men will act alike, nor will all women), but all Twix bars taste the same. Thus, leading explanations for these past findings about consumer choice cannot easily explain our findings in the realm of hiring decisions.

In this paper, we propose and explore another mechanism to account for our findings, which may also influence consumer choices. Specifically, we examine the salience of group-level diversity. Because groups can have emergent properties (e.g., diversity) that are not apparent when considering individuals, people choosing groups may weight these emergent group properties in their decisions. Past research has proposed this mechanism as a contributor to variety-seeking in consumer choice domains (Kahneman & Lovallo, 1993; Read et al., 2001), but it has received scant empirical attention.

Across 6 preregistered experiments (n = 3,509), we provide evidence for the isolated choice effect. Consistent with our theorizing, we find the salience of group diversity mediates the effect and drawing attention to diversity attenuates the effect.

STUDY 1

In Study 1, we examined how the isolated choice effect influences the gender diversity of job candidates chosen in a hypothetical hiring scenario. We predicted that participants tasked with hiring for a single position (i.e., those randomized to an isolated choice condition) would choose a lower proportion of women than participants tasked with hiring for multiple positions (i.e., those randomized to a collective choice condition).

Methods

Participants. We decided in advance to recruit 525 participants through Amazon's Mechanical Turk. After excluding participants who did not follow directions (following our preregistration plan), we were left with 500 participants (48.8% of whom

identified as men). Participants were paid $0.70 to fill out a survey that took about five minutes to complete. This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=4kg79v).

Procedure. Participants were asked to imagine they were hiring for a technology company that was looking to fill five different roles: software engineer, product manager, user experience designer, marketing analyst, and sales representative.

Participants were randomly assigned to either the *isolated choice* condition or the *collective choice* condition. In the *isolated choice* condition, participants were told they would be hiring one person to fill one of these five roles. In the *collective choice* condition, participants were told they would be hiring five people – one person to fill each role. As a result, participants in the *isolated choice* condition made one hiring decision, while participants in the *collective choice* condition made five hiring decisions simultaneously (i.e., all five decisions were shown on the same screen). To balance the number of hiring decisions made across conditions, we assigned five times as many participants to the *isolated choice* condition as the *collective choice* condition.

Before making hiring decisions, participants were shown descriptions of each of the five roles the organization was seeking to fill (e.g., "A software engineer writes computer code to design, develop, maintain, test, and evaluate computer software."). In the *isolated choice* condition, participants were shown the same five job descriptions but were randomly assigned to fill just one of the five roles. After reading these job descriptions,

participants were asked to make hiring decisions for the role(s) they were responsible for filling.

For each role, participants were asked to choose among three candidates who had prior work experience in a relevant job. The candidates were held constant across conditions, so the decisions participants had to make were identical across conditions; all that differed was the number of decisions participants were responsible for making. Because participants in both conditions always chose among three candidates for each position, participants in both conditions were engaging in joint evaluation (as opposed to separate evaluation) for each decision (Bohnet et al., 2015), as will be the case across all studies in this paper. The three candidates for each role always included at least two men, and we varied candidate quality such that the woman always had a moderate amount of experience.[16] In addition, for one role, we included three men to obscure our study's focus on gender diversity. Participants were provided with each candidate's picture (taken from the Chicago Face Database; Ma, Correll, & Wittenbrink, 2015), most recent job, and number of years of experience. All study materials are available in our Online Supplement.

Results

[16]The woman always had fewer years of experience than one of the men and more years of experience than the other (e.g., 4 years versus 5 years or 2 years). We made the woman marginally less qualified than one of the men in order to avoid ceiling effects that might arise from social desirability concerns in experiments. In Studies 2 and 4A, we replicate our effects when there are no differences in quality across candidates by gender.

Our dependent variable of interest was whether a woman was selected in each hiring decision.[17] In the *isolated choice* condition, women were chosen in 7.4% of all hiring decisions; in the *collective choice* condition, women were chosen in 18.0% of all hiring decisions. Following our preregistered analysis plan, we ran an ordinary least squares regression with robust standard errors clustered by participant to predict whether a female candidate was chosen in each hiring decision. Because the unit of analysis was a single hiring decision, each participant in the *isolated choice* condition was included once, while each participant in the *collective choice* condition was included in the regression five times. Our only predictor variable was an indicator variable for being in the *isolated choice* condition. We found that the effect of being in the *isolated choice* condition on the likelihood of selecting a female candidate was significant ($b_{isolated\_choice}$ = -0.105, SE = 0.026, $p < 0.001$; 95% CI: [-0.157, -0.054]).[1819] In other words, making isolated choices produced less gender-diverse groups of hires than making sets of choices.[20]

---

[17] For decisions where all three candidates were men, the dependent variable was coded as zero in both conditions. Our results are identical regardless of whether we include these decisions in our analyses.

[18] As a robustness check, we reran all analyses in all studies using logistic regressions rather than OLS regressions. None of the results change in significance based on the model used (see the Online Supplement, page 4).

[19] Complete regression tables are presented in our Online Supplement for all regressions in this manuscript.

[20] In this study, the decision in which all three candidates were men included one black man. As an exploratory analysis, we tested whether the rate of choosing the black man varied across conditions. Consistent with the isolated choice effect, we found that the

Discussion

In this study, we found that people hired a lower proportion of women when making personnel decisions in isolation rather than in collections. In a supplemental study, we replicated this effect using a different organizational context and different stimuli (see Online Supplement Study S1).

One potential concern about the design of this study, however, is that participants in the *collective choice* condition had to make five times as many decisions as those in the *isolated choice* condition, so fatigue or depletion could be responsible for our findings. To address this concern, we also ran a preregistered replication of this study where participants in the *isolated choice* condition made four additional, unrelated decisions (e.g., choosing between couches) in order to hold the number of decisions constant between conditions. Participants were assigned to one of the five hiring decisions at random, and the hiring decision was presented in the same order across conditions (first, second, third, fourth, or fifth) to balance the timing of decisions across conditions. We again replicated our results (see Online Supplement Study S2), suggesting that our findings cannot be explained by fatigue.

STUDY 2

In Study 2, we ran a conceptual replication of Study 1 using a different study paradigm. Our new paradigm involved a more natural and familiar task with authentic

---

black man was chosen marginally more often in the *collective choice* condition (39.1%) than in the *isolated choice* condition (24.7%), $z = 1.75$, $p = 0.08$.

stimuli: participants were asked to select famous authors for inclusion in a high school English class.

Methods

Participants. We decided in advance to recruit 600 participants through Amazon's Mechanical Turk. After excluding participants who did not follow directions (following our preregistration plan), we were left with 598 participants (53.0% of whom identified as men). This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=6sw3eg).

Procedure. Participants were truthfully told we were interested in understanding which authors Americans think students should be exposed to in high school. They were told they would see a list of 25 authors whose books are commonly read in American high schools and that they would be asked to recommend some number of those authors for inclusion in a high school English course (and we did, in fact, share their recommendations with a high school English teacher who had influence in determining curriculum at their school). Eight of the 25 authors were women. The authors were chosen by combining recommendations from GoodReads and Buzzfeed (Althouse, 2013).

Participants were randomly assigned to either an *isolated choice* condition or a *collective choice* condition and saw a list of the names and photos of the same 25 authors in each condition. In the *isolated choice* condition, participants were asked to recommend one author. They were asked to select this sole author by ranking all of the authors in

their mind and selecting their #1, #2, #3, #4, or #5 ranked author for inclusion in a high school English course. We randomly assigned participants in the *isolated choice* condition to tell us either their #1, #2, #3, #4, or #5 ranked author. In the *collective choice* condition, participants were asked to recommend five authors. They were asked to select this set of authors by ranking all of the authors in their mind and selecting their top five authors for inclusion in a high school English course. Thus, in both conditions, participants were asked to rank all 25 of the authors, but they reported either their top five or one member of their top five at random. Since participants saw all 25 authors in both conditions, all decisions were made under joint evaluation.

This procedure ensured that the aggregation of *isolated* decisions across participants was equivalent in overall quality to the aggregation of *collective* decisions across participants, as both sets of decisions should contain equal numbers of #1, #2, #3, #4, and #5 ranked authors. However, in the *collective choice* condition, diversity should be more salient because participants are asked to report a group of five authors as opposed to only one individual author. All study materials are available in our Online Supplement.

Results

Our outcome of interest was whether participants recommended female authors for inclusion in an American high school English course. In the *isolated choice* condition, 24.0% of the recommended authors were women; in the *collective choice* condition, 29.5% of the recommended authors were women. Following our preregistered analysis

plan, we ran an ordinary least squares regression with robust standard errors clustered by participant to predict the selection of a female author. As in past studies, the unit of analysis was a single author recommendation, so participants in the *collective choice* condition contributed five times as many observations to our regression as participants in the *isolated choice* condition. Our only predictor variable was an indicator variable for being in the *isolated choice* condition. We found that being in the *isolated choice* condition significantly decreased participants' likelihood of selecting a female author ($b_{isolated\_choice}$ = -0.0545, p = 0.047, 95% CI: [-0.108, -0.0008]). This study offers further evidence that isolated choices lead to the selection of less gender-diverse groups than collective choices.

STUDY 3

In Study 3, we tested our proposed mechanism. We explored whether the isolated choice effect arises because diversity is less salient when choices are made in isolation than collectively. Study 3A is a mediation study, while Study 3B is a moderation study.

Study 3A

In Study 3A, we tested whether diversity (a property of groups but not individuals) is more salient when making collective choices than when making isolated choices and whether the salience of diversity mediates people's personnel selection decisions.

Methods

Participants. We decided in advance to recruit 520 participants through Amazon's Mechanical Turk. After excluding participants who did not follow directions (following our preregistration plan), we were left with 502 participants (43.8% of whom identified as men). Participants were paid $0.60 to take a survey that could be completed in about five minutes. This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=q76fa4).

Procedures. We used the same stimuli and a similar study design to Study 1. As in Study 1, participants were asked to imagine they were hiring for a technology company that was looking to fill five different roles. Participants were randomly assigned to either the *isolated choice* condition or the *collective choice* condition. Those in the *isolated choice* condition were told they would be tasked with hiring one person to fill one of these five roles. Those in the *collective choice* condition were told they would be tasked with hiring five people, one for each role.

For each role, participants were asked to choose among three candidates who had prior work experience in a relevant job. The candidates were held constant across conditions. The three candidates for each job always included at least two men (and we included three men as candidates for one job to obscure the fact that our study was focused on gender diversity). Given our focus on gender diversity, in order to avoid wasting participants in this study, we did not assign any participants in the *isolated choice* condition to hire for the role where it was impossible to select a woman, and to maintain parallelism in our design, we discarded hiring decisions made by participants in

the *collective choice* condition for this role in our analyses (as per our preregistration). Participants were provided with each candidate's picture (taken from the Chicago Face Database; Ma, Correll, & Wittenbrink, 2015), most recent job, and number of years of experience.

After participants had made their hiring selection(s), to test for our mechanism, we asked participants to what extent they agreed with the following statement: "I considered how my choice(s) would influence the diversity of the tech team hired when making my decision(s)" on a scale from 1 (Not at all) to 7 (Extremely). All study materials are available in our Online Supplement.

Results

Our dependent variable of interest was whether a woman was selected in each hiring decision. Consistent with our previous results, in the *isolated choice* condition, women were chosen in 15.3% of the hiring decisions; in the *collective choice* condition, women were chosen in 21.1% of the hiring decisions ($b_{isolated\_choice}$ = -0.058, SE = .030, p = 0.054; 95% CI: [-0.117, 0]). Participants also reported that diversity was considered less in their decision-making process in the *isolated choice* condition ($M_{isolated\_choice}$ = 3.32, SD = 2.09) than in the *collective choice* condition ($M_{collective\_choice}$ = 4.03, SD = 2.02; $t(500) = 3.02$, $p = 0.0027$).

We next tested whether considering diversity mediated the relationship between making isolated choices and selecting female candidates. First, there was a significant main effect of assignment to the *isolated choice* condition on how much participants

considered diversity ($b_{isolated\_choice}$ = -0.704, SE = 0.232, p = 0.0023). Second, the relationship between considering diversity and selecting a female candidate was also significant ($b_{considering\_diversity}$ = 0.059, SE = 0.0070, $p < 0.001$). Consistent with mediation, the effect of assignment to the *isolated choice* condition on selecting a female candidate ($b_{isolated\_choice}$ = -0.061, SE = 0.030, p = 0.040)[21] was eliminated when controlling for diversity considerations ($b_{isolated\_choice}$ = -0.020, SE = 0.028, p = 0.49). A Sobel test confirmed that this reduction in effect size was significant ($b_{reduction}$ = -0.042, SE = 0.015, p = 0.0046), and a 5,000-sample bootstrap analysis (MacKinnon, Fairchild, & Fritz, 2007; Shrout & Bolger, 2002) also produced a bias-corrected 95% confidence interval for the size of the indirect effect that excluded zero (95% CI: [-0.073, -0.015]).

Study 3B

Given the inherent limitations of mediation analyses, we also tested our mechanism through a moderation study in which we manipulated rather than measured whether diversity was salient. Specifically, in Study 3B, we tested whether the isolated choice effect is eliminated when attention is drawn to diversity in both the isolated and collective choice conditions.

Methods

Participants. We decided in advance to recruit 1,050 participants through Amazon's Mechanical Turk. After excluding participants who did not follow directions

[21] Due to bootstrapping standard errors in the mediation analysis, estimates from the mediation analysis vary slightly from estimates from the main regression.

(following our preregistration plan), we were left with 1,038 participants (44.1% of whom identified as men). Participants were paid $0.45 to take a survey that could be completed in about four minutes. This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=5sx59i).

Procedures. Similar to Studies 1 and 3A, participants in this study were asked to imagine they were hiring for a technology company that was looking to fill five different roles. This study used a 2x2 (*isolated choice* vs. *collective choice* x *diversity valued* vs. *control*) factorial design. Participants were randomly assigned to either the *isolated choice* condition or the *collective choice* condition. To balance the number of hiring decisions made across conditions, we assigned four times as many participants to the *isolated choice* condition as the *collective choice* condition. Those in the *isolated choice* condition were told they would be tasked with hiring one person to fill one of these five roles. Those in the *collective choice* condition were told they would be tasked with hiring five people, one for each role.

Participants were also randomly assigned to either a condition where attention was drawn to diversity by stating that it was valued or a control condition where no such statement was made. In the *diversity valued* condition, participants were told "The company strongly values diversity" immediately prior to making their hiring selection(s); in the *control* condition, we omitted this statement, so diversity was not made explicitly salient.

Participants had to choose among three candidates for each role they were asked to fill. These candidates were held constant across conditions and always included at least two men (one of the decisions included three men as candidates to conceal that our study focused on gender diversity). Participants were provided with each candidate's picture (taken from the Chicago Face Database; Ma, Correll, & Wittenbrink, 2015), most recent job, and number of years of experience. Finally, as a manipulation check, we asked participants to rate their agreement with the statement "The company strongly values diversity" on a scale from "1: Strongly disagree" to "5: Strongly agree."

All study materials are available in our Online Supplement.

Results

Our manipulation check confirmed that participants in the *diversity valued* condition believed that the organization valued diversity more than those in the *control* condition ($M_{diversity\_valued} = 4.42$, $SD_{diversity\_valued} = 0.90$; $M_{control} = 3.24$, $SD_{control} = 0.67$; $t(1,036) = 23.85$, $p < 0.001$), suggesting that our manipulation was successful.

Our dependent variable of interest was whether a woman was selected in each hiring decision. When attention was not drawn explicitly to diversity, we found evidence of the isolated choice effect: participants in the *isolated choice* condition hired women 15.0% of the time, and participants in the *collective choice* condition hired women 25.3% of the time ($b_{isolated\_choice} = -0.103$, $SE = .030$, $p < 0.001$; 95% CI: [-0.163, -0.044]; see Figure 1). However, when attention was drawn to diversity by telling participants the organization valued diversity, the isolated choice effect disappeared: participants in the *isolated choice*

condition hired women 36.5% of the time, and participants in the *collective choice* condition hired women 37.1% of the time ($b_{isolated\_choice}$ = -0.006, *SE* = .034, *p* = 0.867; 95% CI: [-0.072, 0.060]; see Figure 1). There was also a significant interaction between the *diversity valued* and *isolated choice* conditions ($b_{isolated\_choice*diversity\_valued}$ = 0.10, *SE* = .045, *p* = 0.03; 95% CI: [0.009, 0.187]), suggesting drawing attention to diversity moderated the effect of being in the *isolated choice* condition on the likelihood of selecting a female candidate.

Discussion

Taken together, Studies 3A and 3B provide evidence that isolated choices lead to less diverse hires because diversity is less salient when choices are made in isolation than when they are made in collections. Study 3A shows that participants attend less to diversity when making isolated choices than collective choices, and this mediates the effect of isolated choices on the gender diversity of hired candidates. Study 3B shows that the isolated choice effect is eliminated when attention is drawn to diversity considerations by explicitly noting that a company values diversity.

In a supplemental study, we ran a 2x2 experiment where we varied whether choices were made in isolation or collectively as well as whether we told participants the organization they were helping had low levels of gender diversity (19% of its employees were women) or high levels of gender diversity (48% of its employees were women). When gender diversity was high (and therefore there would be no specific reason to attend to diversity), we replicated the isolated choice effect, but when gender diversity was low (and diversity

was thus a salient problem), we no longer observed a significant isolated choice effect (see Online Supplement Study S3). This supplementary study provides further evidence consistent with our proposed mechanism: diversity is more salient in collective choices than isolated choices, which leads to more diverse hires when choices are made collectively.

STUDY 4

In Study 4, we examined whether the isolated choice effect extends to real–rather than hypothetical–decisions.

Study 4A

In Study 4A, participants were tasked with hiring graduate students to participate in a business pitch competition. Their bonuses were contingent upon the success of the candidate(s) they hired.

Methods

Participants. We decided in advance to recruit 310 participants through Amazon's Mechanical Turk. After excluding participants who did not follow directions (following our preregistration plan), we were left with 271 participants (49.1% of whom identified as men). Participants were paid $0.40, plus a potential bonus of up to $1.00, to take a survey that took about five minutes to complete. This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=pq3wz5).

Procedure. Participants were truthfully told that an East Coast university would be hosting a business pitch competition. People would compete in teams, and each team

would include three members: a team leader, a financial analyst, and a brand manager. Each team would come up with a business idea, and judges would evaluate these business ideas to choose a winning team.

Participants were randomly assigned to either an *isolated choice* or *collective choice* condition. Participants in the *isolated choice* condition were told they would hire one person to join a pitch competition team, and if the team that person joined won the pitch competition, the participant would earn a $1.00 bonus. Participants in the *collective choice* condition were told that they would hire all three members of a pitch competition team, and if their team won the pitch competition, the participant would earn a $1.00 bonus. After the conclusion of the experiment, we organized and ran the pitch competition as described, and participants were paid bonuses accordingly.

All candidates for the pitch competition team were actual graduate students at the East Coast university where the pitch competition took place. Participants were provided with fully truthful information about the candidates including photos, names, the number of years they had completed in their graduate programs, and their areas of research. Candidates for the team leader role included two white men and one white women with comparable backgrounds. Candidates for the financial analyst and brand manager roles were three white men. While participants in the *collective choice* condition hired one person for each of the three roles, we did not assign any participants in the *isolated choice condition* to make hiring decisions for the financial analyst or brand manager roles because we were interested in whether a woman was hired for each position, and all three

candidates for those roles were men. All study materials are available in our Online Supplement.

Results

Our dependent variable of interest was whether the woman was hired for the team leader role. As preregistered, we conducted a two-sample test of proportions to compare the rate of choosing the female candidate across conditions. Consistent with our previous studies, we found that the woman was hired significantly less often in the *isolated choice* condition (20.6%) than in the *collective choice* condition (45.7%), $z = 4.26$, $p < 0.001$.

Study 4B

In Study 4B, participants were tasked with recommending speakers for an academic conference, and their choices helped determine who was invited to the conference.

Methods

Participants. We decided in advance to recruit 600 participants through Amazon's Mechanical Turk (53.5% of whom identified as men). Participants were paid $0.45 to take a survey that took about four minutes to complete. This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=t4qh4s).

Procedure. Participants were told that the researchers conducting this study were organizing an upcoming academic conference. They were told that they would recommend speakers for the conference, and the conference would include at least one speaker from each of five focus areas: Health, Education, Conflict Management,

Financial Literacy, and Energy. We further informed participants that their decisions were consequential because we would invite the most frequently recommended speakers from each focus area to the conference.

Participants were randomly assigned to either the *isolated choice* condition or the *collective choice* condition. In the *isolated choice* condition, participants were tasked with choosing one speaker from one of the focus areas chosen at random (either Health, Education, Conflict Management, Financial Literacy or Energy). In the *collective choice* condition, participants were tasked with selecting five speakers, one from each focus area.

For each focus area, participants were asked to choose among three candidates. These candidates were held constant across conditions. For four of the five focus areas, the three candidates included two white men and one white woman. To obfuscate the study's focus on gender diversity, we included three white men as candidates for one focus area. Because we were interested in whether a woman was selected for each decision, participants in the *isolated choice* condition were never assigned to choose a speaker in the focus area with three men as candidates, and we did not include decisions made by participants in the *collective choice* condition in this focus area in our analyses (to ensure the conditions were evenly balanced). To balance the number of hiring decisions made across conditions, we assigned four times as many participants to the *isolated choice* condition as the *collective choice* condition.

Participants were provided with truthful information about all candidates including pictures, names, academic institutions, years of academic experience, and impact factors (their h-index on Google scholar, as of August 2019). All study materials are available in our Online Supplement.

Results

Our dependent variable of interest was whether a woman was chosen for each decision. In the *isolated choice* condition, women were chosen in 32.2% of all selection decisions; in the *collective choice* condition, women were chosen in 45.8% of all selection decisions. Following our preregistered analysis plan, we ran an ordinary least squares regression with robust standard errors clustered by participant where the dependent variable was whether a woman was chosen in each decision. The only predictor variable was an indicator variable for being in the *isolated choice* condition. We found that the effect of being in the *isolated choice* condition on the likelihood of selecting a female candidate was significant and negative ($b_{isolated\_choice}$ = -0.163, SE = 0.034, $p < 0.001$; 95% CI: [-0.230, -0.096]). In other words, consistent with our other studies, making isolated choices produced less gender-diverse groups of speakers for the conference than making sets of choices.

GENERAL DISCUSSION

Across six preregistered experiments, we find that the isolated choice effect influences the gender composition of groups. We present evidence that people select less gender-diverse candidates when making isolated hiring or selection decisions (i.e., when

making a single hire) than when making collections of selection decisions (i.e., when making multiple hires). We also find that diversity is less salient when people make isolated choices than collections of choices, and salience of diversity mediates the effects of isolated choices on personnel selection decisions. Together, our results highlight a potentially important contributing factor to the underrepresentation of women in many groups and organizations, given that hires are often made in isolation rather than in collections.

All of the studies presented explore settings where women are underrepresented, and women are hired infrequently in the isolated choices we study. In settings where men are underrepresented, our theory predicts that hiring in collections (rather than in isolation) should still increase the gender diversity of hired candidates, but increasing gender diversity in these settings would mean increasing the rate of selecting *male* candidates. To test this prediction, we conducted a supplemental study using the same stimuli as Study 1 but switching the genders of all job candidates. In other words, there were more qualified women than men available to hire for each position. We still found that people opted for less gender diversity when making isolated choices than when making sets of choices (see Online Supplement Study S4); however, because these revised stimuli included an overrepresentation of qualified women, participants hired fewer *men* when making isolated choices as opposed to fewer women. These results provide support for the idea that the isolated choice effect is about diversity and not just women.

We chose to study the isolated choice effect in the context of gender diversity because of its important policy implications. However, the isolated choice effect should generalize to other contexts where group diversity is considered desirable. It would be valuable for future research to examine how our findings extend to other social categories (e.g., race). It would also be useful to test the effects of isolated choices on personnel selection decisions in the field to establish the external validity of these findings. In particular, future field work comparing hiring decisions made on separate occasions (that are truly separated in time) with decisions made collectively would be of great value.

Examining the multiple potential motives that underlie people's greater desire for diversity when it is made salient would also be useful. The salience of diversity may affect hiring decisions in multiple ways: people may believe that diversity is better for group performance; they may believe that it is their moral obligation to pursue diversity; or they may want to avoid appearing discriminatory. Disentangling whether all or some of these factors drive the patterns we detect would be beneficial. Study 3B establishes that organizational context, including explicit demands for diversity, can influence the strength of the isolated choice effect. Future research exploring this further would be valuable.

Past research has found that people often react negatively to explicit attempts to increase diversity (Dobbin, Schrage, & Kalev, 2015; Legault, Gutsell, & Inzlicht, 2011; Plant & Devine, 2001) and that it is challenging to change people's biases and stereotypes (Chang et al., 2019; Forscher et al., 2019; Kalev, Dobbin, & Kelly, 2006; Lai

et al., 2016). Prompting people to make collective rather than isolated decisions is a novel approach to increasing diversity in that it does not involve reprimands or explicit directives (which can be viewed as overly paternalistic), nor does it rely on changing people's biases and stereotypes. For these reasons, it may be a particularly promising approach to increasing diversity in organizations.

Practically, our results suggest that organizations interested in increasing diversity might consider having decision makers hire in collective rather than isolated ways. For example, rather than hiring one person every month, a company could hire three people every quarter. In the long run, the company will hire the same number of people, but choices will be made collectively (every quarter) rather than in isolation (every month). Alternatively, companies could give certain people oversight over many hiring decisions so at least some employees are making collective decisions, rather than allowing hiring managers or teams make decisions in isolation.

While our work prescribes structural changes to hiring practices that are likely to increase organizational diversity, we recognize that these prescriptions may be difficult to implement in some contexts. It would therefore be valuable for future work to explore ways of making isolated choices *feel* collective. For example, future research could examine the effects on hiring of showing people their past hiring decisions or the employees most recently hired by their co-workers. Such interventions could also draw attention to diversity, which is a group-level property, and may lead to decisions that look more like collective, rather than isolated, choices.

REFERENCES

Althouse, S. (2013, July 5). 23 Books You Didn't Read In High School But Actually Should. Retrieved November 7, 2018, from BuzzFeed website: https://www.buzzfeed.com/spenceralthouse/23-books-you-didnt-read-in-high-school-but-actually-should

Babcock, L., Recalde, M. P., Vesterlund, L., & Weingart, L. (2017). Gender differences in accepting and receiving requests for tasks with low promotability. *The American Economic Review*, *107*(3), 714–747.

Bohnet, I., Van Geen, A., & Bazerman, M. (2015). When performance trumps gender bias: Joint vs. Separate evaluation. *Management Science*, *62*(5), 1225–1234.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2012). Salience theory of choice under risk. *The Quarterly Journal of Economics*, *127*(3), 1243–1285.

Bordalo, P., Gennaioli, N., & Shleifer, A. (2013). Salience and consumer choice. *Journal of Political Economy*, *121*(5), 803–843.

Camerer, C., Babcock, L., Loewenstein, G., & Thaler, R. (1997). Labor supply of New York City cabdrivers: One day at a time. *The Quarterly Journal of Economics*, *112*(2), 407–441.

Chang, E. H., Milkman, K. L., Gromet, D. M., Rebele, R. W., Massey, C., Duckworth, A. L., & Grant, A. M. (2019). The mixed effects of online diversity training. *Proceedings of the National Academy of Sciences*, *116*(16), 7778–7783.

Dobbin, F., Schrage, D., & Kalev, A. (2015). Rage against the iron cage: The varied effects of bureaucratic personnel reforms on diversity. *American Sociological Review*, *80*(5), 1014–1044.

Forscher, P. S., Lai, C. K., Axt, J. R., Ebersole, C. R., Herman, M., Devine, P. G., & Nosek, B. A. (2019). A meta-analysis of procedures to change implicit measures. *Journal of Personality and Social Psychology*.

Harrison, D. A., & Klein, K. J. (2007). What's the difference? Diversity constructs as separation, variety, or disparity in organizations. *Academy of Management Review*, *32*(4), 1199–1228.

Jones, S., & Donnelly, G. (2017, June 16). Why We Logged Every Fortune 500 Company's Diversity Data, Or Lack Thereof. *Fortune*. Retrieved from http://fortune.com/2017/06/16/why-we-logged-every-fortune-500-companys-diversity-data-or-lack-thereof/

Kahneman, D., & Lovallo, D. (1993). Timid choices and bold forecasts: A cognitive perspective on risk taking. *Management Science*, *39*(1), 17–31.

Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, *71*(4), 589–617.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., … Marshburn, C. K. (2016). Reducing implicit racial preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001.

Legault, L., Gutsell, J. N., & Inzlicht, M. (2011). Ironic effects of antiprejudice messages: How motivational interventions can reduce (but also increase) prejudice. *Psychological Science*, *22*(12), 1472–1477.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135.

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology*, *58*(1), 593–614. https://doi.org/10.1146/annurev.psych.58.110405.085542

Phillips, L. T., Slepian, M. L., & Hughes, B. L. (2018). *Perceiving groups: The people perception of diversity and hierarchy.*

Plant, E. A., & Devine, P. G. (2001). Responses to other-imposed pro-Black pressure: Acceptance or backlash? *Journal of Experimental Social Psychology*, *37*(6), 486–501.

Read, D., Antonides, G., Van den Ouden, L., & Trienekens, H. (2001). Which is better: Simultaneous or sequential choice? *Organizational Behavior and Human Decision Processes*, *84*(1), 54–70.

Read, D., & Loewenstein, G. (1995). Diversification bias: Explaining the discrepancy in variety seeking between combined and separated choices. *Journal of Experimental Psychology: Applied*, *1*(1), 34.

Read, D., Loewenstein, G., & Rabin, M. (1999). Choice bracketing. *Journal of Risk and Uncertainty*, *19*, 171–197.

Schroeder, J., & Risen, J. L. (2016). Befriending the enemy: Outgroup friendship longitudinally predicts intergroup attitudes in a coexistence program for Israelis and Palestinians. *Group Processes & Intergroup Relations*, *19*(1), 72–93.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and non-experimental studies: New procedures and recommendations. *Psychological Methods*, 422–445.

Simonson, I. (1990). The effect of purchase quantity and timing on variety-seeking behavior. *Journal of Marketing Research*, 150–162.

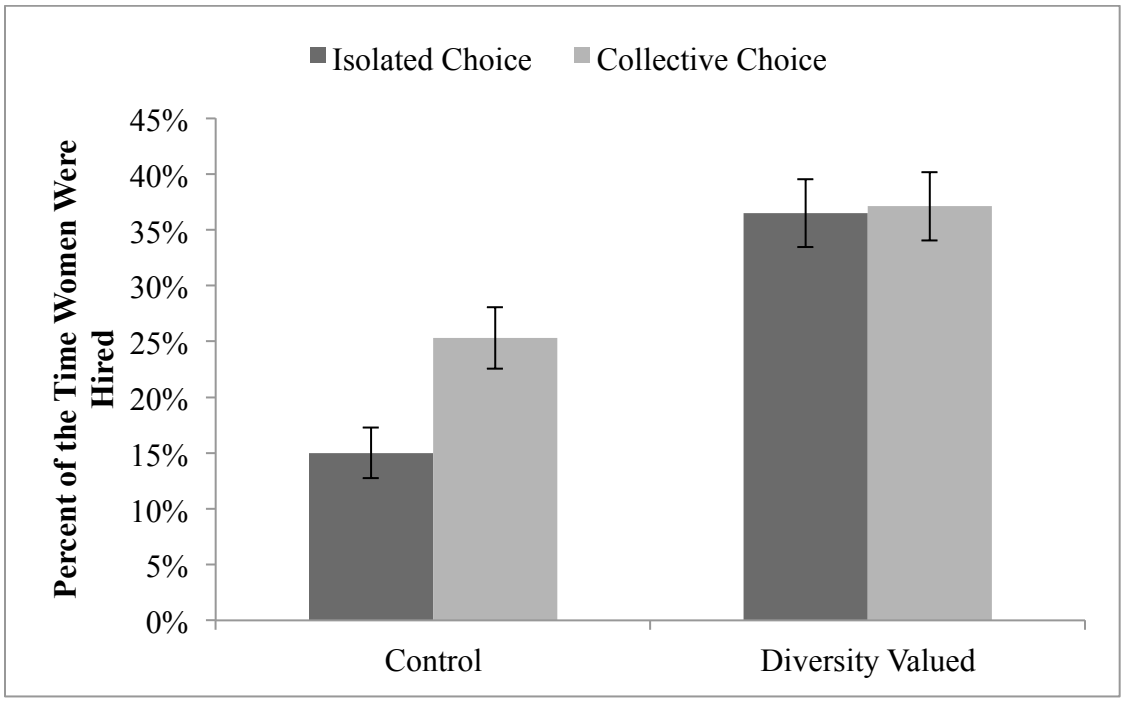
Simonson, I., & Winer, R. S. (1992). The influence of purchase quantity and display format on consumer preference for variety. *Journal of Consumer Research*, *19*(1), 133–138.

Thaler, R. H. (1999). Mental accounting matters. *Journal of Behavioral Decision Making*, *12*(3), 183–206.

FIGURES

Figure 1

Explicitly Stating that Diversity is Valued Moderates the Isolated Choice Effect



CHAPTER 3. "DIVERSITY WASHING": AN ORGANIZATIONAL IMPRESSION MANAGEMENT STRATEGY THAT CREATES THE APPEARANCE OF DIVERSITY

Edward H. Chang

ABSTRACT

Extant theorizing on diversity would suggest that organizations have strong incentives to distance themselves from diversity, given discrimination against women and racial minorities and risks of losing status by associating with lower status groups in society. But in recent years, there has been a marked increase in negative scrutiny applied to organizations that are deemed to be lacking in diversity, creating strong external incentives for organizations to attend to diversity. I propose that to resolve these competing demands, organizations engage in "diversity washing"–whereby they mislead outsiders about their actual diversity levels or practices–to create the appearance of caring about diversity or being more diverse than they are in reality. I provide evidence that organizations' signals of their diversity are indeed distorted. For example, in multiple organizational contexts, employees from demographically underrepresented groups are portrayed in higher proportions in diversity signals than exist in reality in organizations. Diversity washing seems to be exacerbated when the signals are more visible, consistent with it being driven by impression management concerns. This work highlights the potential strength of impression management motives in guiding diversity-related

decisions in organizations and suggests that research on diversity should consider impression management as a key motivation in theorizing and scholarship.

INTRODUCTION

Copious research has demonstrated that people have negative stereotypes and biases about members of historically marginalized groups (e.g., women, racial minorities; Fiske et al., 2002; Rosette et al., 2008). These biases can translate into discrimination against members of these groups, as women and racial minorities often do not receive the same opportunities to succeed in organizations as do men and White people (Bertrand & Mullainathan, 2004; Kang et al., 2016; Milkman et al., 2015; Moss-Racusin et al., 2012). As a result, many organizations and professions remain relatively homogeneous, as women and racial minorities face barriers to entry (Bertrand & Mullainathan, 2004; Moss-Racusin et al., 2012).

But in recent years, there has been a rapid increase in scrutiny on the topic of diversity, and in particular, negative scrutiny directed at organizations that are deemed to have inadequate diversity. For example, multiple countries have passed laws mandating minimum levels of gender diversity on the boards of public companies in those countries (Miller, 2014); institutional investors are placing demands on companies to add more diversity or risk losing shareholder support (Broughton, 2019); and entire industries are facing pressures to hire more women and underrepresented racial minorities (Marcus, 2015). This suggests that even if organizations do not want diversity, they may have external incentives to care about diversity in order to avoid penalties and negative attention for lacking diversity.

How do these competing demands around diversity get resolved in organizations?

I propose that organizations may engage in "diversity washing" (i.e., greenwashing for diversity; Delmas & Burbano, 2011; Laufer, 2003) whereby they deliberately attempt to mislead outsiders about their actual diversity levels or practices. For example, companies could distort signals of their diversity by overrepresenting people from underrepresented groups in externally visible settings like websites (e.g., which employees are featured on company websites) or panels (e.g., who companies send to conferences or job fairs) relative to their actual proportions within the companies, or companies could tout how much they care about diversity, in spite of not being demographically diverse. Diversity washing can help organizations resolve competing demands as they get to project an image of diversity to appease outsiders who may scrutinize them, but they do not actually have to be diverse in reality.

In documenting and exploring the phenomenon of diversity washing, this paper helps provide an updated understanding of what motivates diversity-related decisions in organizations. Prevailing theories about diversity and people's treatment of members of historically underrepresented groups would predict that organizations should, if anything, distance themselves from diversity given ubiquitous bias and discrimination against women and racial minorities. This stand in contrast to greenwashing–a phenomenon where companies mislead consumers about their environmentally performance (Delmas & Burbano, 2011)–as there is not much theorizing or evidence that being perceived as environmentally friendly would be aversive to companies (cf. Kim & Lyon, 2014). However, extant theories on diversity fail to account for the new reality of increasingly

high external demands for diversity, which create strong impression management incentives around diversity. More generally, incorporating impression management into our understanding of diversity helps theorizing make more accurate predictions about contemporary diversity-related decisions in organizations, including explaining why phenomena like diversity washing occur.

THEORY AND HYPOTHESES

Reasons for Organizations to Hide Diversity

Past research and theorizing provides voluminous evidence as to why organizations may want to hide diversity within their ranks or otherwise distance themselves from demographic diversity. For example, research has documented people's persistent negative biases and stereotypes about women and racial minorities (Berdahl & Min, 2012; Fiske et al., 2002; Rosette et al., 2008; Rudman & Glick, 2001), field experiments have shown that people discriminate against women and racial minorities in a variety of contexts (Bertrand & Mullainathan, 2004; Milkman et al., 2015; Moss-Racusin et al., 2012; Neumark et al., 1996), and research has shown that people penalize female and racial minority leaders and the organizations they lead (Brooks et al., 2014; Heilman, 2001; Kanze et al., 2018; Rosette et al., 2008). All of this research suggests that organizations would not want to and do not hire members of historically underrepresented groups at the same rates as they do members of majority groups, leading to a lack of diversity in organizations.

Even if organizations already have diversity among their ranks, they may not

want to showcase this diversity to outsiders. Organizations typically prefer to be higher in status rather than lower in status (Sauder et al., 2012). Given that associating with low status others can lead to decreases in status (Podolny, 1993) and that women and racial minorities are historically low status groups in society (Bonam et al., 2016; Castilla, 2008; Fiske et al., 2002; Levanon et al., 2009; Murphy & Oesch, 2016; Rosette et al., 2008), organizations should try to disassociate themselves from women and racial minorities and not embrace demographic diversity in order to preserve status. Indeed, research has found that occupations that become increasingly female-dominated are subsequently devalued and lose status (Levanon et al., 2009; Murphy & Oesch, 2016), the physical spaces associated with Black Americans are negatively stereotyped and perceived as lower in status (Bonam et al., 2016), and investors penalize companies that appoint women to boards because these companies are perceived to be embracing diversity (Solal & Snellman, 2019). Thus, in order to project or maintain status, it may not be good for organizations to showcase associations with low status groups in society, suggesting that they should not want to highlight associations with women and racial minorities.

Beyond the discrimination that individuals from historically underrepresented groups face, diverse teams are also subject to negative evaluator biases. People associate diversity in teams with relationship conflict, which creates biased perceptions of the amount of relationship conflict in diverse as opposed to homogeneous teams and lowers people's willingness to support diverse teams (Lount Jr et al., 2015). A recent meta-

analysis also showed that diverse teams receive lower outsider evaluations of their performance as compared to objective performance benchmarks, suggesting that people negatively view diverse teams (Van Dijk et al., 2012).

Finally, diversity is also threatening to majority group members in society (e.g., White people) and may make it harder for organizations to recruit majority group members. Past research has found that White people find multiculturalism threatening (Plaut et al., 2011), and White people think they will be discriminated against at companies that espouse pro-diversity values (Dover et al., 2016). Further, reminders of increasing racial diversity in society can lead majority racial group members to express more negative racial attitudes and increase intergroup hostility (Craig & Richeson, 2014b, 2014a; Danbold & Huo, 2015), which can affect their willingness to join organizations with dissimilar others. In other words, showcasing diversity may make it harder for organizations to recruit talent in the job market, which can harm their long-term competitiveness and viability.

Together, extant research on diversity suggests many reasons why companies might not want to showcase demographic diversity, and in fact, indicates that they have incentives to downplay diversity. Highlighting demographic diversity in an organization would seem to harm how others perceive it, as past research and theory would suggest that organizations that are perceived as more diverse would be seen as lower in status, would be evaluated more negatively, and would have a harder time recruiting people from majority groups in society. Thus, one might make the following prediction about

how organizations will choose to distort signals of their diversity:

*Hypothesis 1: Organizations will not choose to highlight demographic diversity. In other words, people from demographically underrepresented groups will be portrayed in **lower** proportions than exist in reality in diversity signals..*

Impression Management as a Motivator to Attend to Diversity

In opposition to the research documenting reasons why organizations may want to distance themselves from diversity, there also exists research positing reasons why organizations may want to strive for diversity. Scholars have suggested that diversity can provide benefits to team and organizational performance, and embracing diversity can be seen as a moral imperative given past discrimination against members of historically underrepresented groups (Ely & Thomas, 2001). For example, correlational research suggests that companies with higher levels of racial diversity and gender diversity perform better (Dezsö & Ross, 2012; Hunt et al., 2015; Richard, 2000), and research suggests that diverse groups perform better than homogeneous ones on a variety of different dimensions (Cox et al., 1991; Gaither et al., 2017; Hoogendoorn et al., 2013; Keck & Tang, 2017; Lount Jr & Phillips, 2007; Loyd et al., 2013; McLeod et al., 1996; Sommers, 2006; Woolley et al., 2010). In addition, given the importance of justice and equity in organizations (Colquitt et al., 2001; Greenberg & Colquitt, 2013), diversity can be seen as evidence of a fair and just process where all people have equal opportunity and access to organizations, regardless of their social category or background.

These reasons in favor of diversity may lead organizations to be targeted for

lacking diversity. Indeed, in recent years, negative scrutiny towards organizations that are deemed to have inadequate diversity has dramatically increased. For example, some countries have passed laws mandating minimum levels of gender diversity on corporate boards of public companies (Miller, 2014; Smale & Miller, 2015); law firms have come under intense scrutiny for announcing partner classes lacking sufficient gender and racial diversity (Scheiber & Eligon, 2019); and Forbes magazine had to publicly acknowledge that they "blew it" after receiving backlash for publishing a list of 100 innovative leaders that included only one woman (McGregor, 2019).

To avoid this negative scrutiny around lacking diversity, organizations may engage in impression management behaviors around diversity. Organizations typically want to avoid negative scrutiny because it can have negative consequences for their reputations (Desai, 2011). As a result, they engage in impression management around scrutinized behaviors to try to escape critical attention (Bolino et al., 2008; Elsbach et al., 1998). Recent research confirms that some organizations may attend to diversity for impression management reasons to avoid negative scrutiny from sources like the media. For example, research has found that organizations may pay lip service to diversity by using pro-diversity language in job advertisements but not follow through on those values in practice (Kang et al., 2016), and organizations appear to strive to reach minimum thresholds for gender diversity on corporate boards that are perceived to help them escape negative scrutiny (Chang, Milkman, Chugh, et al., 2019).

Beyond impression management motives to escape negative scrutiny,

organizations could also have impression management motives to attend to diversity in efforts to attract more diverse job candidates if they also believe in the organizational benefits or moral imperatives of diversity (Avery & McKay, 2006). Homophily predicts that people want to be a part of groups and organizations that are composed of people who look like them (McPherson & Smith-Lovin, 1987; McPherson et al., 2001). If an organization lacks people from a particular demographic category and believes that homophily will attract people from that demographic category, the organization may distort signals about its diversity levels and overrepresent people from that demographic category. Past research provides some evidence that portraying an image of people from particular demographic categories may be an effective strategy for attracting people from those demographic categories. For example, people from marginalized racial groups in the United States tend to be attracted to organizations that represent them (e.g., when a company's website depicts Black employees and managers, Black people are more attracted to that company; Avery, 2003; Avery, Hernandez, & Hebl, 2004), and past research has found that diversity cues are important for attracting members of racial minorities (Purdie-Vaughns et al., 2008).

As a result of these motivations to attend to diversity for impression management reasons, organizations may engage in "diversity washing"–or attempts to create an image of having a diverse workforce–as a way to avoid negative scrutiny and/or increase workforce diversity. For both motivations, organizations should want to distort diversity signals in ways that depict people from demographically underrepresented groups in

higher proportions than exist in reality (e.g., portraying the organization as having more Black employees than the organization has in reality if Black employees are underrepresented in the organization relative to society), as these strategies should project an image of being more diverse. If we believe these impression management motives are strong enough, then we might make the following prediction about how organizations will distort signals of their diversity:

Hypothesis 2: Organizations will engage in "diversity washing." In other words, people from demographically underrepresented groups will be portrayed in ***higher*** *proportions than exist in reality in diversity signals.*

Moderating Effects of Visibility

If it is true that organizations engage in diversity washing (as opposed to downplaying diversity), then we might expect diversity washing behavior to be moderated by the visibility of the group or organization. Visibility refers to how much attention groups or organizations receive, regardless of reason (Chiu & Sharfman, 2011). When organizations are more visible, they tend to face greater external pressures and therefore have greater incentives to attend to impression management behaviors of all sorts (Chiu & Sharfman, 2011; Gardberg & Fombrun, 2006), including diversity washing. For example, past research has shown that more visible organizations are more likely to conform to social norms around diversity behaviors for impression management reasons (Chang, Milkman, Chugh, et al., 2019), and that organizations that are more visible engage in more prosocial actions for impression management reasons after being

the target of social movement boycotts (McDonnell & King, 2013).

As a result, increased visibility should mean that organizations are more likely to engage in diversity washing as compared to less visible groups or organizations. Increased visibility translates to greater scrutiny and greater pressures to engage in impression management behaviors like diversity washing, and failing to meet expectations around issues like diversity can lead to greater punishments highly visible firms. Thus, assuming diversity washing occurs, then we might make the following prediction about the circumstances that would exacerbate the behavior:

Hypothesis 3: Increased visibility will lead to increased diversity washing.

Overview of Studies

In Studies 1 and 2, I test whether members of demographically underrepresented groups are portrayed in systematically higher or lower proportions than exist in reality in signals of organizational diversity. In Study 1, I examine diversity signals of technology companies, and in Study 2, I examine diversity signals of law firms. In both contexts, I do not find evidence that organizations downplay diversity; in fact, I find evidence of diversity washing, whereby employees from demographically underrepresented groups–namely Black people and women–are portrayed in systematically higher proportions on technology company and law firm websites than exist in reality. In Study 3, I experimentally manipulate visibility to see if increased visibility exacerbates diversity washing, as would be predicted by an impression management account. Finally, in Study 4, I experimentally manipulate visibility using a different paradigm, again to see if

increased visibility exacerbates diversity washing.

STUDY 1: DIVERSITY SIGNALS ON TECHNOLOGY COMPANY WEBSITES

In Study 1, I analyzed data from EEO-1 reports and company job or career websites to test whether diversity is downplayed or whether diversity washing occurs on the job or career websites of technology companies. I compared demographic compositions of employees featured on websites to demographic compositions of actual employees at these companies to see whether there are systematic distortions of proportions of employees from demographically underrepresented groups.

Methods

Data. The EEO-1 data I used were compiled by the Center for Investigative Reporting (Evans & Rangarajan, 2017). The Center for Investigative Reporting sought to obtain EEO-1 reports for 2016 from 211 of the largest technology companies in the San Francisco Bay Area. They compiled the list of 211 companies by taking the top 150 publicly traded technology companies as ranked by worldwide annual revenue and including 61 private companies that had valuations of at least $1 billion, as estimated by Crunchbase and CB Insights. Twenty-two companies released EEO-1 report data to Reveal, and I used all of these data in my analyses.

EEO-1 reports are mandatory government reports for companies with 100 employees or more (*EEO-1: Answers to Filing Questions Often Asked by Employers*,

n.d.). They include data on the racial and gender composition[22] of all employees at a company, broken down by job category (e.g., executive/senior officials and managers). I paired these EEO-1 data with data collected from these same companies' job or career websites. For each company, I had a mixed-race and mixed-gender group of research assistants go to the homepage of the company and navigate to find the landing page for their job or career website. The research assistants then worked together as a group to code the race and gender of all employees depicted on the job or career website using the same categories used in the EEO-1 reports. While there are undoubtedly issues with the methodology of categorizing people based on appearance rather than asking for self-identified racial and gender identity, past research suggests that people can accurately identify the self-identified race of the subject of a photograph with high accuracy (Blascovich et al., 1997; Chance & Goldstein, 1981). In addition, given the focus of the study on signals of diversity, perceived racial and gender categories may be more important than self-identified racial and gender categories.

Analysis Strategy. To test whether and how signals of diversity are distorted on company websites, I compared the demographic compositions of the employees depicted on their job or career websites to the demographic compositions of their employees as reported to the federal government via EEO-1 reports. Specifically, I used Monte Carlo simulations to generate null distributions under the assumption that employees were

---

[22] Gender is categorized as male or female. Race is categorized as American Indian or Alaskan Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, Two or More Races, or White.

chosen at random to be depicted on the job or career website of a company. Thus, for each company and each demographic category (e.g., Black people at Apple), I calculated how likely or unlikely it was (i.e., with p-values) that that proportion of people from a given demographic category (or a more extreme proportion) would appear on the job or career website of the company if the company were selecting from its employees at random.

Beyond calculating these likelihoods for individual companies and demographic categories, I also considered all 22 companies simultaneously to see whether people from certain demographic categories were systematically overrepresented or underrepresented as a whole across companies. Using techniques from meta-analysis, I combined p-values across companies to test the probability of seeing these levels of distortions (or more extreme levels) across companies, assuming the null hypothesis that companies randomly selected employees for their websites.

Results

Summary Statistics. The companies in my dataset varied greatly in size, with the smallest company (23andMe) having 297 employees and the largest (Apple) having 77,192 employees. On average, these companies were composed of 29.8% Asian Americans, 4.0% Black or African Americans, 5.8% Hispanic or Latinx Americans, and 57.5% White Americans. On average, 33.4% of employees at these companies were women.

The companies also varied in terms of the number of employees they depicted on their job or career websites, ranging from one employee (View) to 44 employees (Twitter), with an average of 16.9 employees depicted. On average, of the employees depicted on these websites, 31.7% were Asian Americans, 14.1% were Black or African Americans, 3.6% were Hispanic or Latinx Americans, and 48.5% were White Americans. On average, 47.9% of the employees depicted on these websites were women. These summary statistics alone suggests some degree of distortion of diversity signals.

How Are Signals of Diversity Distorted for Technology Companies? In contrast to Hypothesis 1 (which states that companies will downplay diversity) and consistent with Hypothesis 2 (which states that companies will engage in diversity washing), Black people and women of all races were portrayed in systematically higher proportions on technology company job or career websites relative to their proportions among employees at these companies. Sixteen out of 22 companies (72.7%) had a higher proportion of Black employees on their job or career websites relative to their actual proportion among their employees (overall probability across companies: $p < 0.001$; see Figure 1), and 18 out of 22 companies (81.8%) had a higher proportion of women (overall probability across companies: $p < 0.001$; see Figure 2).

Insert Figure 1 about here

Insert Figure 2 about here

STUDY 2: DIVERSITY SIGNALS ON LAW FIRM WEBSITES

In Study 2, I replicated the analyses of technology companies in a different organizational context–law firms. I compared data from the Law Firm Diversity Survey distributed by Vault and the Minority Corporate Counsel Association about attorney demographics at different law firms with data taken from their job or career websites to see whether and how diversity signals of law firms are distorted.

Methods

Data. I examined the largest 150 law firms in the United States according to PublicLegal as of 2019. I combined this list of law firms with the 2020 Vault Law 100, a national ranking of the most prestigious law firms based on the assessments of lawyers at peer firms. Given that there was significant overlap between these two lists, these two lists gave me a total of 162 firms.

For each of these firms, I searched the Law Firm Diversity Survey database to see if the firm had a corresponding survey response in 2019. These surveys report the demographics of attorneys at law firms broken down by job title, gender, and race.[23] I aggregated attorney demographics by summing the categories "Associate", "Equity

---

[23] Gender is categorized as man or woman. Race is categorized as White/Caucasian, African-American/Black, Hispanic/Latinx, Alaska Native/American Indian, Asian, Native Hawaiian/Pacific Islander, or Multiracial.

Partner", "Non-Equity Partner", and "Of Counsel." 142 of the 162 law firms had submitted a survey response in 2019.

I paired these demographic data with data collected from these same law firms' job or career websites using the same methodology as in Study 1 of coding the gender and race of employees featured on the career websites of these law firms. Unlike the technology companies in Study 1, many law firms do not depict any employees on their websites, leaving a final sample of 95 law firms (i.e., law firms that both featured at least one person on their website and submitted a survey response in 2019).[24]

Results

Summary Statistics. The law firms in my dataset ranged from 147 attorneys to 2,116 attorneys. On average, attorneys were 8.2% Asian, 3.4% Black, 4.1% Hispanic/Latinx, and 82.0% White at these law firms. On average, 36.7% of attorneys at these law firms were women.

The law firms also varied in terms of the number of employees they depict on their job or career websites, ranging from one employee to 26 employees, with an average of 6.4 employees depicted. On average, of the employees depicted on these websites, 14.3% were Asian, 17.6% were Black, 4.3% were Latinx, and 61.7% were

[24] Whether a law firm chooses to depict employees on its website does not appear related to observable firm characteristics like status or size, but it is negative correlated with gender diversity. In other words, law firms with higher proportions of female attorneys are less likely to depict any employees on their websites ($p = 0.038$).

White. On average, 53.9% of the employees depicted on these websites were women. These summary statistics again suggests some degree of distortion of diversity signals.

How Are Signals of Diversity Distorted for Law Firms? I used the same Monte Carlo simulation method as in Study 1 to see whether signals of diversity are distorted on law firm websites. Failing to support Hypothesis 1 and providing additional support for Hypothesis 2, which states that companies will engage in diversity washing, Black people and women of all races were systematically portrayed in higher proportions on law firm job or career websites relative to their actual proportions among employees at these companies. Sixty out of 95 companies (63.2%) had a higher proportion of Black employees on their job or career websites relative to their actual proportion among their employees (overall probability across companies: $p < 0.001$), and 71 out of 95 companies (74.7%) had a higher proportion of women of all races (overall probability across companies: $p < 0.001$).

Discussion of Studies 1 and 2

In Studies 1 and 2, I tested whether and how companies distort signals of their diversity. While much of past research on diversity would predict that organizations should, if anything, downplay signals of their demographic diversity (Hypothesis 1), I find no evidence of this. On the contrary, I find evidence that members of demographically underrepresented groups were portrayed in systematically higher proportions on technology company and law firm websites than exist in actuality among their employees, a phenomenon I label "diversity washing." These results are consistent

with Hypothesis 2, which states that impression management motives surrounding diversity will lead companies to portray higher proportions of members of demographically underrepresented groups in signals of diversity than they have in reality.

Both technology and law are industries where demographic diversity is lacking relative to the overall U.S. population. For example, in spite of Black people composing roughly 13% of the U.S. population and women composing roughly 50% of the population (*US Census: Quick Facts*, 2016), on average, only 4.0% and 3.4% of technology and law firm employees were Black, respectively, in the companies in my dataset, while only 33.4% and 36.7% of employees were women. The fact that these industries are relatively homogeneous compared to the overall U.S. population could contribute to why these companies engage in diversity washing as opposed to downplaying their diversity. As a result, it is unclear whether these findings would generalize to industries or companies that have higher proportions of demographic diversity, much less industries or companies with higher levels of demographic diversity than the overall U.S. population.

STUDY 3: INCREASED VISIBILITY EXACERBATES DIVERSITY WASHING IN AN EXPERIMENT

Having established diversity washing as a phenomenon in at least two organizational contexts, in Study 3, I sought to test whether diversity washing becomes more severe as the visibility of a group increases, as would be suggested by impression management motives. To test this, I ran an online experiment where I tested people's

preferences for diversity for an externally visible group (e.g., the employees featured in a newsletter sent to external stakeholders) versus an internal group (e.g., the employees featured in a newsletter sent to internal stakeholders). First, diversity washing predicts that people should be more likely to choose Black employees and women for either group relative to their base rate availability (i.e., people will overrepresent Black employees and women; Hypothesis 2). Second, people should be more likely to choose Black employees and women for the externally visible group than the internal group (i.e., increasing visibility increases diversity washing; Hypothesis 3).

Method

Participants. Four hundred and two participants in the U.S. were recruited through Amazon's Mechanical Turk to participate in a short online research study (57.0% identified as men). These participants were paid $0.35 for completing a survey they were told would take 2-3 minutes of their time. This study was preregistered on AsPredicted.org (http://aspredicted.org/blind.php?x=x4ge6q).

Procedures. Participants were asked to imagine they worked for a small technology company of roughly twenty employees. They were then shown pictures of twenty faces taken from the Chicago Face Database (Ma et al., 2015). The twenty faces were composed of one Black man, one Latinx man, four Asian men, two Asian women, eight White men, and four White women, as these proportions roughly reflect the average demographic compositions of the technology companies in my sample in Study 1. The Chicago Face Database provides data on how often people categorize each face as male

or female and as Asian, Black, Latinx, White, Multiracial, or Other. I only selected faces where pretests had shown that people had at least 90% accuracy in categorizing gender and at least 90% accuracy in categorizing race (e.g., for the Latinx man, over 90% of people correctly identified him as a man, and over 90% of people correctly identified him as Latinx). I also stimulus sampled the faces to ensure the results were not driven by the particular faces chosen.

To manipulate visibility, participants were randomly assigned to either the *externally visible group* condition or the *internal group* condition. In the *externally visible group* condition, participants were told they were tasked with choosing five employees to feature in a newsletter that would be sent to all shareholders of the company (i.e., a group of external stakeholders where impression management concerns may apply). In the *internal group* condition, participants were told they were tasked with choosing five employees to feature in a newsletter that would be sent to all employees of the company (i.e., a group of internal stakeholders where impression management concerns should be less likely to apply since all employees presumably already know the underlying actual diversity of the company). Participants then chose five employees to feature in the newsletter.

Finally, as a manipulation check, participants were asked to what extent they agreed with the statement, "The group of employees I chose for the newsletter will be scrutinized by the public or media," on a scale from 1 (Strongly Disagree) to 7 (Strongly Agree).

Results and Discussion

First, I tested whether people chose the Black and women employees for the groups significantly more often than would be expected by chance, as would be predicted by diversity washing (Hypothesis 2). If people were choosing employees at random, 5% of the employees chosen for the groups should be Black (since 1 out of 20 of the employees was Black), while 30% of the employees chosen should be women (since 6 out of 20 of the employees were women). Across conditions, 14.8% of the employees chosen were Black, which was significantly greater than expected, $t(401) = 22.28$, $p < 0.001$, and 39.7% of the employees chosen were women, again significantly greater than expected, $t(401) = 15.12$, $p < 0.001$. These results are consistent with the results from Studies 1 and 2 and with Hypothesis 2, which states that people will portray higher proportions of members of demographically underrepresented groups in diversity signals than they have in reality.

Second, I tested whether manipulating visibility exacerbated the extent of diversity washing. The manipulation check showed that people felt that the group would be under greater scrutiny in the *externally visible group* condition ($M = 4.86$, $SD = 2.05$) as compared to the *internal group* condition ($M = 2.18$, $SD = 1.61$), $t(400) = 14.6$, $p < 0.001$, suggesting that I successfully manipulated visibility and thus impression management concerns. Next, the Black employee represented 15.8% of group members chosen in the *externally visible group* condition and only 13.8% of groups members in the *internal group* condition, $t(400) = 2.34$, $p = 0.020$, $d = 0.23$. Put another way, 79.0%

of participants chose to include the Black employee in their group in the *externally visible group* condition, while only 68.8% of participants chose to include the Black employee in their group in the *internal group* condition. In other words, increasing the visibility of the group increased the rate at which people engaged in diversity washing of Black people, consistent with Hypothesis 3.

With regards to diversity washing of women, women represented 39.6% of group members chosen in the *externally visible group* condition and 39.8% of group members chosen in the *internal group* condition, $t(400) = 0.157$, $p = 0.875$. In this case, increasing visibility did not exacerbate the diversity washing of women, failing to provide additional support for Hypothesis 3.

STUDY 4: MANIPULATING VISIBILITY USING A DIFFERENT PARADIGM

In Study 4, I again sought to test whether diversity washing becomes more severe as the visibility of a group increases, but I used a different paradigm to manipulate visibility and different stimuli. In this study, I manipulated visibility of a chosen group by telling participants that the photos and names of group members would be featured on a website or that the photos and names of group members would not be featured on a website. Diversity washing predicts that people should be more likely to choose Black people and women for either group relative to their base rate availability (Hypothesis 2). In addition, people should be more likely to choose Black people and women when group members' photos and names are featured on a website than when group members' photos and names are not featured (Hypothesis 3).

Method

Participants. Four hundred U.S. participants were recruited on Prolific to participate in a short online research study (40.0% identified as men). These participants were paid $0.35 for completing a survey they were told would take 2-3 minutes of their time. This study was preregistered on AsPredicted.org (https://aspredicted.org/blind.php?x=xf9f9c).

Procedures. Participants were told that a research center was looking to highlight on its website some of the research that scientists affiliated with the center were doing. They were told the website would feature four stories from scientists and that we were looking for input on who should be featured. To manipulate visibility, participants were randomly assigned to either the *identifiable information* condition or the *anonymous* condition. In the *identifiable information* condition, participants were told that each story from a scientist would be accompanied by that scientist's name and photo. In the *anonymous* condition, participants were told that each story from a scientist would not have a name or photo accompany it. If participants were worried about diversity for impression management reasons, these concerns should have been more salient in the *identifiable information* condition, as the (lack of) demographic diversity of the stories on the website would be more visible than in the *anonymous* condition.

Participants were then shown pictures, names, and areas of research (e.g., education) of twelve real social scientists affiliated with an actual research center. The twelve scientists were composed of one Black man, one Asian man, six White men, one

Asian woman, and three White women. Participants then chose four of these scientists to recommend be featured on the website.

Results and Discussion

First, I tested whether participants chose the Black and women scientists to be featured on the website significantly more often than would be expected by chance, as would be predicted by diversity washing (Hypothesis 2). If participants were choosing scientists at random, 8.33% of the scientists chosen for the website should be Black (since 1 out of 12 of the scientists was Black), while 33.33% of the scientists chosen should be women (since 4 out of 12 of the scientists were women). Across conditions, 10.6% of the scientists chosen were Black, which was significantly greater than expected, $t(399) = 3.61$, $p < 0.001$, and 39.6% of the scientists chosen were women, again significantly greater than expected, $t(399) = 6.72$, $p < 0.001$. These results provide support for Hypothesis 2, which states that people will portray higher proportions of members of demographically underrepresented groups in diversity signals than exist in reality.

Second, I tested whether manipulating visibility exacerbated the extent of diversity washing. The Black scientist represented 11.7% of group members chosen in the *identifiable information* condition and only 9.5% of groups members in the *anonymous* condition. This difference in rates of diversity washing between conditions was in the expected direction, but it was only marginally significant, $t(398) = 1.83$, $p = 0.0675$. Similarly, women represented 41.2% of group members chosen in the

identifiable information condition and only 38.0% of groups members in the *anonymous* condition. Again, this difference was in the expected direction, but it was only marginally significant, $t(398) = 1.76$, $p = 0.0795$. These results provide marginal support that increasing visibility exacerbates diversity washing (Hypothesis 3).

GENERAL DISCUSSION

In this paper, I document the phenomenon of diversity washing. In multiple organizational contexts and in experiments, I find evidence that Black people and women of all races are systematically overrepresented in diversity signals, such that they appear in higher proportions than actually exist in organizations. Diversity washing occurs in spite of copious past research on diversity which would suggest that organizations should be wary of showcasing diversity. Organizations may be driven to diversity wash for impression management reasons, as there now exist external incentives for companies to attend to diversity to avoid negative scrutiny. I also find some evidence that diversity washing is exacerbated by higher visibility, consistent with past research showing that visibility leads to higher incentives to engage in impression management (Chang, Milkman, Chugh, et al., 2019; Chiu & Sharfman, 2011; McDonnell & King, 2013).

This work contributes to the relatively limited literature examining the importance of impression management as a motivator of diversity-related decisions and behaviors. Past research has documented how impression management motives may shape how individuals react in interracial settings (Apfelbaum et al., 2008) and may influence diversity-related hiring decisions (Chang, Milkman, Chugh, et al., 2019). But the

distortions of diversity signals due to diversity washing I document can be completely detached from actual diversity levels within companies, which distinguishes diversity washing from actual changes in diversity-related hiring decisions documented in past research (Chang, Milkman, Chugh, et al., 2019). Diversity washing should instead be seen as a form of organizational decoupling akin to greenwashing (Delmas & Burbano, 2011; Laufer, 2003), where an organization's behaviors do not align with what it says about its behaviors (in the case of greenwashing, with regards to its environmental record or performance). Contrary to greenwashing, which assumes that the only real risks to greenwashing for organizations are if the deception is discovered (cf. Kim & Lyon, 2014), diversity washing represents a much more risky proposition to organizations, given theory and research which suggests that associating with diversity can damage an organization's standing (in addition to the risk of the deception being discovered). Thus, diversity washing represents a more extreme phenomenon in which impression management concerns may dominate other motives and suggests that impression management should be considered as a primary motivator of attention to diversity in organizations.

While I document systematic overrepresentations of members of certain demographically underrepresented groups (e.g., Black people, women of all races) across companies, the analysis strategy I use does not allow me to disentangle the motivations of any individual company. The results I find are consistent with impression management motives surrounding diversity, but they may not actually reflect strategic or deliberate

decision making. For example, people may have heuristics that they should choose to represent at least one person from every demographic group in external signals of diversity, similar to how some investors use naïve diversification strategies and invest in equal amounts across all funds offered by their retirement plans (Benartzi & Thaler, 2001). If people and organizations were using such a heuristic, we might also expect to see overrepresentation of Latinx employees in diversity signals, given that Latinx people represent the largest minority group in the U.S. and are underrepresented in both technology companies and law firms relative to the general U.S. population. However, I find no evidence of overrepresentation of Latinx people in any of my studies.

If the diversity washing I document is driven by impression management, questions remain as to whether diversity washing is driven primarily by desires to avoid negative scrutiny or by desires to attract a more diverse workforce. If organizations are engaging in diversity washing to attract diversity, however, it may be a bad strategy. Past research has found that diversity statements–when they are found not to be backed up by action–can actually decrease organizational attraction among job candidates because people dislike inauthenticity (Windscheid et al., 2016). In addition, diversity washing may inflate perceptions of organizational dishonesty around diversity, which can make organizations less attractive to racial minorities (Wilton et al., 2019). In other words, organizations may be engaging in an impression mismanagement strategy (Steinmetz et al., 2017) if they diversity wash as a means to increase their attraction to women and racial minorities.

An important limitation of the organizational contexts I examine empirically (i.e., technology companies and law firms) is they are contexts in which women and racial minorities are underrepresented relative to the general population. Organizations may feel pressure to portray themselves as having diversity that reflects the general population (Ely & Thomas, 2001), which may drive diversity washing in organizations where women and racial minorities are underrepresented relative to the general population, but not in organizations where women and racial minorities are proportionally represented or overrepresented. Future work should examine how organizations behave in contexts where women and racial minorities are proportionally represented or overrepresented relative to the general population. In these contexts, organizations may have other impression management strategies around diversity and may not engage in diversity washing.

Technology and law are also industries that are high status overall (e.g., white collar work; typically high paying). Higher status firms may face greater pressures to diversity wash for a variety of reasons. First, higher status firms are typically more visible than lower status firms (Chiu & Sharfman, 2011; McDonnell & King, 2013), which would suggest that they have greater incentives to engage in impression management behaviors like diversity washing. Second, people typically have higher expectations for the behaviors of higher status firms (McDonnell & King, 2018; Rhee & Haunschild, 2006). To the extent that diversity is seen as an expectation of firms, higher status firms should also be more wary of failing to diversity wash given that they face greater

penalties for failing to meet higher expectations (McDonnell & King, 2018; Rhee & Haunschild, 2006). Future research should explore how impression management behaviors around diversity play out in lower status firms and industries.

In addition, in my experiments, I find only limited evidence that experimentally manipulating visibility increases diversity washing. In part, this may be driven by the fact that even in the less visible conditions, participants may have strong impression management or social desirability concerns around diversity, as they may worry about appearing sexist or racist to experimenters even in anonymous surveys (Apfelbaum et al., 2008; Kawakami et al., 2009). This is evidenced by the large amounts of diversity washing that participants exhibit in these less visible conditions, suggesting that there may not be much further room to push people's impression management concerns and thus their diversity washing behaviors in artificial online experiments.

More generally, in augmenting our understanding of what drives diversity-related decisions in organizations, this work potentially provides guidance about how we might increase diversity in organizations. Much of past research on interventions to increase diversity in organizations has focused on attempting to change people's biases and stereotypes (e.g., through diversity training). Unfortunately, past research has shown that many of these interventions are not particularly successful (Chang, Milkman, Gromet, et al., 2019; Kalev et al., 2006; Lai et al., 2016). In highlighting the role of impression management, this work illuminates other factors and motivations that may influence diversity-related decisions in organizations. Ways to make impression management

concerns around diversity more salient in hiring decisions may thus be a fruitful way to increase diversity in selection decisions.

REFERENCES

Apfelbaum, E. P., Sommers, S. R., & Norton, M. I. (2008). Seeing race and seeming racist? Evaluating strategic colorblindness in social interaction. *Journal of Personality and Social Psychology*, *95*(4), 918.

Avery, D. R. (2003). Reactions to diversity in recruitment advertising—Are differences black and white? *Journal of Applied Psychology*, *88*(4), 672.

Avery, D. R., Hernandez, M., & Hebl, M. R. (2004). Who's Watching the Race? Racial Salience in Recruitment Advertising1. *Journal of Applied Social Psychology*, *34*(1), 146–161.

Avery, D. R., & McKay, P. F. (2006). Target practice: An organizational impression management approach to attracting minority and female job applicants. *Personnel Psychology*, *59*(1), 157–187.

Benartzi, S., & Thaler, R. H. (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review*, *91*(1), 79–98.

Berdahl, J. L., & Min, J.-A. (2012). Prescriptive stereotypes and workplace consequences for East Asians in North America. *Cultural Diversity and Ethnic Minority Psychology*, *18*(2), 141.

Bertrand, M., & Mullainathan, S. (2004). Are Emily and Greg more employable than Lakisha and Jamal? A field experiment on labor market discrimination. *The American Economic Review*, *94*(4), 991–1013.

Blascovich, J., Wyer, N. A., Swart, L. A., & Kibler, J. L. (1997). Racism and racial categorization. *Journal of Personality and Social Psychology*, *72*(6), 1364.

Bolino, M. C., Kacmar, K. M., Turnley, W. H., & Gilstrap, J. B. (2008). A multi-level review of impression management motives and behaviors. *Journal of Management*, *34*(6), 1080–1109.

Bonam, C. M., Bergsieker, H. B., & Eberhardt, J. L. (2016). Polluting Black Space. *Journal of Experimental Psychology: General*, *145*(11), 1561.

Brooks, A. W., Huang, L., Kearney, S. W., & Murray, F. E. (2014). Investors prefer entrepreneurial ventures pitched by attractive men. *Proceedings of the National Academy of Sciences*, *111*(12), 4427–4431.

Broughton, K. (2019, February 13). Small Companies Face Added Pressure to Put Women on Boards. *Wall Street Journal*. https://www.wsj.com/articles/small-companies-face-added-pressure-to-put-women-on-boards-11550070322

Castilla, E. (2008). Gender, Race, and Meritocracy in Organizational Careers. *American Journal of Sociology*, *113*(6), 1479–1526.

Chance, J. E., & Goldstein, A. G. (1981). Depth of processing in response to own-and other-race faces. *Personality and Social Psychology Bulletin*, *7*(3), 475–480.

Chang, E. H., Milkman, K. L., Chugh, D., & Akinola, M. (2019). Diversity Thresholds: How Social Norms, Visibility, and Scrutiny Relate to Group Composition. *Academy of Management Journal*, *62*(1), 144–171.

Chang, E. H., Milkman, K. L., Gromet, D. M., Rebele, R. W., Massey, C., Duckworth, A. L., & Grant, A. M. (2019). The mixed effects of online diversity training. *Proceedings of the National Academy of Sciences*, *116*(16), 7778–7783.

Chiu, S.-C., & Sharfman, M. (2011). Legitimacy, visibility, and the antecedents of corporate social performance: An investigation of the instrumental perspective. *Journal of Management*, *37*(6), 1558–1585.

Colquitt, J. A., Conlon, D. E., Wesson, M. J., Porter, C. O., & Ng, K. Y. (2001). Justice at the millennium: A meta-analytic review of 25 years of organizational justice research. *Journal of Applied Psychology*, *86*(3), 425.

Cox, T. H., Lobel, S. A., & McLeod, P. L. (1991). Effects of ethnic group cultural differences on cooperative and competitive behavior on a group task. *Academy of Management Journal*, *34*(4), 827–847.

Craig, M. A., & Richeson, J. A. (2014a). More diverse yet less tolerant? How the increasingly diverse racial landscape affects white Americans' racial attitudes. *Personality and Social Psychology Bulletin*, *40*(6), 750–761.

Craig, M. A., & Richeson, J. A. (2014b). On the precipice of a "majority-minority" America: Perceived status threat from the racial demographic shift affects White Americans' political ideology. *Psychological Science*, *25*(6), 1189–1197.

Danbold, F., & Huo, Y. J. (2015). No longer "All-American"? Whites' defensive reactions to their numerical decline. *Social Psychological and Personality Science*, *6*(2), 210–218.

Delmas, M. A., & Burbano, V. C. (2011). The drivers of greenwashing. *California Management Review*, *54*(1), 64–87.

Desai, V. M. (2011). Mass media and massive failures: Determining organizational efforts to defend field legitimacy following crises. *Academy of Management Journal*, *54*(2), 263–278.

Dezsö, C. L., & Ross, D. G. (2012). Does female representation in top management improve firm performance? A panel data investigation. *Strategic Management Journal*, *33*(9), 1072–1089.

Dover, T. L., Major, B., & Kaiser, C. R. (2016). Members of high-status groups are threatened by pro-diversity organizational messages. *Journal of Experimental Social Psychology*, *62*, 58–67.

EEO-1: Answers to Filing Questions Often Asked by Employers. (n.d.). Retrieved March 26, 2019, from https://www.eeoc.gov/employers/eeo1survey/faq.cfm

Elsbach, K. D., Sutton, R. I., & Principe, K. E. (1998). Averting expected challenges through anticipatory impression management: A study of hospital billing. *Organization Science*, *9*(1), 68–86.

Ely, R. J., & Thomas, D. A. (2001). Cultural diversity at work: The effects of diversity perspectives on work group processes and outcomes. *Administrative Science Quarterly*, *46*(2), 229–273.

Evans, W., & Rangarajan, S. (2017, October 19). Hidden figures: How Silicon Valley keeps diversity data secret. *Reveal*. https://www.revealnews.org/article/hidden-figures-how-silicon-valley-keeps-diversity-data-secret/

Fiske, S. T., Cuddy, A. J., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, *82*(6), 878.

Gaither, S. E., Apfelbaum, E. P., Birnbaum, H. J., Babbitt, L. G., & Sommers, S. R. (2017). Mere Membership in Racially Diverse Groups Reduces Conformity. *Social Psychological and Personality Science*, 1948550617708013.

Gardberg, N. A., & Fombrun, C. J. (2006). Corporate citizenship: Creating intangible assets across institutional environments. *Academy of Management Review*, *31*(2), 329–346.

Greenberg, J., & Colquitt, J. A. (2013). *Handbook of organizational justice*. Psychology Press.

Heilman, M. E. (2001). Description and prescription: How gender stereotypes prevent women's ascent up the organizational ladder. *Journal of Social Issues*, *57*(4), 657–674.

Hoogendoorn, S., Oosterbeek, H., & Van Praag, M. (2013). The impact of gender diversity on the performance of business teams: Evidence from a field experiment. *Management Science*, *59*(7), 1514–1528.

Hunt, V., Layton, D., & Prince, S. (2015). Diversity matters. *McKinsey & Company*, *1*, 15–29.

Kalev, A., Dobbin, F., & Kelly, E. (2006). Best practices or best guesses? Assessing the efficacy of corporate affirmative action and diversity policies. *American Sociological Review*, *71*(4), 589–617.

Kang, S. K., DeCelles, K. A., Tilcsik, A., & Jun, S. (2016). Whitened resumes: Race and self-presentation in the labor market. *Administrative Science Quarterly*, *61*(3), 469–502.

Kanze, D., Huang, L., Conley, M. A., & Higgins, E. T. (2018). We ask men to win and women not to lose: Closing the gender gap in startup funding. *Academy of Management Journal*, *61*(2), 586–614.

Kawakami, K., Dunn, E., Karmali, F., & Dovidio, J. F. (2009). Mispredicting affective and behavioral responses to racism. *Science*, *323*(5911), 276–278.

Keck, S., & Tang, W. (2017). Gender composition and group confidence judgment: The perils of all-male groups. *Management Science*.

Kim, E.-H., & Lyon, T. P. (2014). Greenwash vs. Brownwash: Exaggeration and undue modesty in corporate sustainability disclosure. *Organization Science*, *26*(3), 705–723.

Lai, C. K., Skinner, A. L., Cooley, E., Murrar, S., Brauer, M., Devos, T., Calanchini, J., Xiao, Y. J., Pedram, C., & Marshburn, C. K. (2016). Reducing implicit racial

preferences: II. Intervention effectiveness across time. *Journal of Experimental Psychology: General*, *145*(8), 1001.

Laufer, W. S. (2003). Social accountability and corporate greenwashing. *Journal of Business Ethics*, *43*(3), 253–261.

Levanon, A., England, P., & Allison, P. (2009). Occupational feminization and pay: Assessing causal dynamics using 1950–2000 US census data. *Social Forces*, *88*(2), 865–891.

Lount Jr, R. B., & Phillips, K. W. (2007). Working harder with the out-group: The impact of social category diversity on motivation gains. *Organizational Behavior and Human Decision Processes*, *103*(2), 214–224.

Lount Jr, R. B., Sheldon, O. J., Rink, F., & Phillips, K. W. (2015). Biased perceptions of racially diverse teams and their consequences for resource support. *Organization Science*, *26*(5), 1351–1364.

Loyd, D. L., Wang, C. S., Phillips, K. W., & Lount Jr, R. B. (2013). Social category diversity promotes premeeting elaboration: The role of relationship focus. *Organization Science*, *24*(3), 757–772.

Ma, D. S., Correll, J., & Wittenbrink, B. (2015). The Chicago face database: A free stimulus set of faces and norming data. *Behavior Research Methods*, *47*(4), 1122–1135.

Marcus, B. (2015, August 12). The Lack Of Diversity In Tech Is A Cultural Issue. *Forbes*. https://www.forbes.com/sites/bonniemarcus/2015/08/12/the-lack-of-diversity-in-tech-is-a-cultural-issue/#751c5e8179a2

McDonnell, M.-H., & King, B. (2013). Keeping up appearances reputational threat and impression management after social movement boycotts. *Administrative Science Quarterly*, *58*(3), 387–419.

McDonnell, M.-H., & King, B. G. (2018). Order in the court: How firm status and reputation shape the outcomes of employment discrimination suits. *American Sociological Review*, *83*(1), 61–87.

McGregor, J. (2019, September 10). 'We blew it': Forbes named 99 men and only one woman on its list of 'most innovative leaders.' *Washington Post*. https://www.washingtonpost.com/business/2019/09/10/we-blew-it-forbes-named-men-only-one-woman-its-list-most-innovative-leaders/

McLeod, P. L., Lobel, S. A., & Cox Jr, T. H. (1996). Ethnic diversity and creativity in small groups. *Small Group Research*, *27*(2), 248–264.

McPherson, J. M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, 370–379.

McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, *27*(1), 415–444.

Milkman, K. L., Akinola, M., & Chugh, D. (2015). What happens before? A field experiment exploring how pay and representation differentially shape bias on the pathway into organizations. *Journal of Applied Psychology*, *100*(6), 1678.

Miller, C. C. (2014, June 19). Women on the Board: Quotas Have Limited Success. *The New York Times*. https://www.nytimes.com/2014/06/20/upshot/women-on-the-board-quotas-have-limited-success.html

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences*, *109*(41), 16474–16479.

Murphy, E., & Oesch, D. (2016). The feminization of occupations and change in wages: A panel analysis of Britain, Germany, and Switzerland. *Social Forces*, *94*(3), 1221–1255.

Neumark, D., Bank, R. J., & Van Nort, K. D. (1996). Sex discrimination in restaurant hiring: An audit study. *The Quarterly Journal of Economics*, *111*(3), 915–941.

Plaut, V. C., Garnett, F. G., Buffardi, L. E., & Sanchez-Burks, J. (2011). "What about me?" Perceptions of exclusion and Whites' reactions to multiculturalism. *Journal of Personality and Social Psychology*, *101*(2), 337.

Podolny, J. M. (1993). A status-based model of market competition. *American Journal of Sociology*, *98*(4), 829–872.

Purdie-Vaughns, V., Steele, C. M., Davies, P. G., Ditlmann, R., & Crosby, J. R. (2008). Social identity contingencies: How diversity cues signal threat or safety for

African Americans in mainstream institutions. *Journal of Personality and Social Psychology*, *94*(4), 615.

Rhee, M., & Haunschild, P. R. (2006). The liability of good reputation: A study of product recalls in the US automobile industry. *Organization Science*, *17*(1), 101–117.

Richard, O. C. (2000). Racial diversity, business strategy, and firm performance: A resource-based view. *Academy of Management Journal*, *43*(2), 164–177.

Rosette, A. S., Leonardelli, G. J., & Phillips, K. W. (2008). The White standard: Racial bias in leader categorization. *Journal of Applied Psychology*, *93*(4), 758.

Rudman, L. A., & Glick, P. (2001). Prescriptive gender stereotypes and backlash toward agentic women. *Journal of Social Issues*, *57*(4), 743–762.

Sauder, M., Lynn, F., & Podolny, J. M. (2012). Status: Insights from organizational sociology. *Annual Review of Sociology*, *38*, 267–283.

Scheiber, N., & Eligon, J. (2019, January 27). Elite Law Firm's All-White Partner Class Stirs Debate on Diversity. *The New York Times*. https://www.nytimes.com/2019/01/27/us/paul-weiss-partner-diversity-law-firm.html

Smale, A., & Miller, C. C. (2015, March 6). Germany Sets Gender Quota in Boardrooms. *The New York Times*. https://www.nytimes.com/2015/03/07/world/europe/german-law-requires-more-women-on-corporate-boards.html

Solal, I., & Snellman, K. (2019). Women Don't Mean Business? Gender Penalty in Board Composition. *Organization Science*, *30*(6), 1270–1288.

Sommers, S. R. (2006). On racial diversity and group decision making: Identifying multiple effects of racial composition on jury deliberations. *Journal of Personality and Social Psychology*, *90*(4), 597.

Steinmetz, J., Sezer, O., & Sedikides, C. (2017). Impression mismanagement: People as inept self-presenters. *Social and Personality Psychology Compass*, *11*(6), e12321.

US Census: Quick Facts. (2016). US Census. https://www.census.gov/quickfacts/

Van Dijk, H., Van Engen, M. L., & Van Knippenberg, D. (2012). Defying conventional wisdom: A meta-analytical examination of the differences between demographic and job-related diversity relationships with performance. *Organizational Behavior and Human Decision Processes*, *119*(1), 38–53.

Wilton, L. S., Bell, A. N., Vahradyan, M., & Kaiser, C. R. (2019). Show Don't Tell: Diversity Dishonesty Harms Racial/Ethnic Minorities at Work. *Personality and Social Psychology Bulletin*, 0146167219897149.

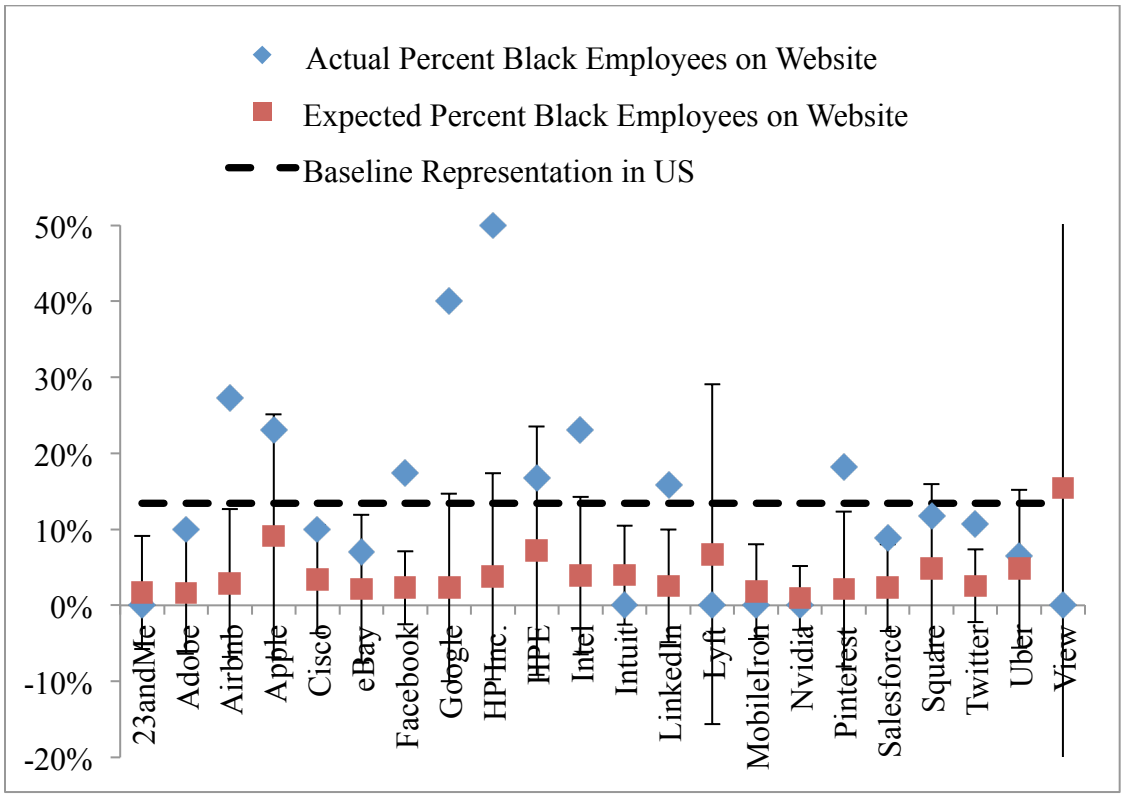
Windscheid, L., Bowes-Sperry, L., Kidder, D. L., Cheung, H. K., Morner, M., & Lievens, F. (2016). Actions speak louder than words: Outsiders' perceptions of diversity mixed messages. *Journal of Applied Psychology*, *101*(9), 1329.

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686–688.

FIGURES

Figure 1

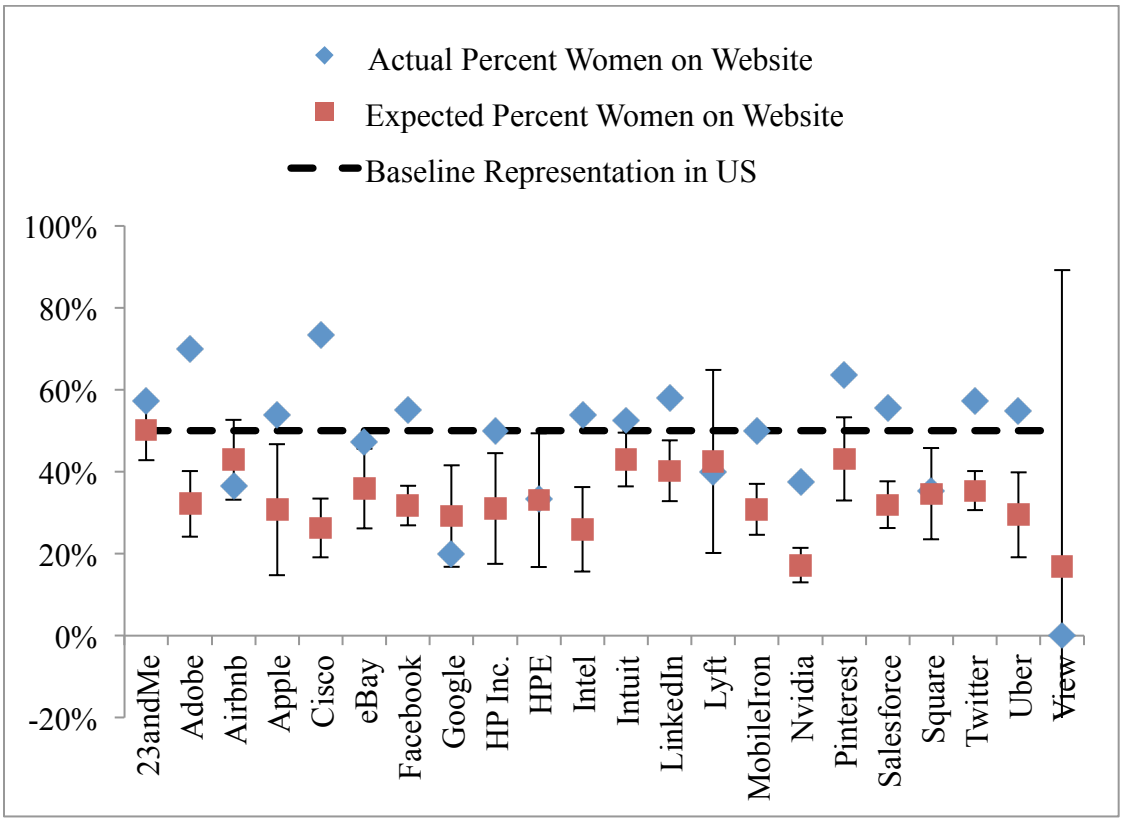
Comparison of Actual Proportion of Black Employees on Job or Career Websites to Expected Proportion Based on Employee Demographics



Error bars represent 95% confidence intervals.

Figure 2

Comparison of Actual Proportion of Women on Job or Career Websites to Expected Proportion Based on Employee Demographics



Error bars represent 95% confidence intervals.